Chapter

Upgrading Non-conventional Yeasts into Valuable Biofactories

Kevin Castillo-Mendieta, Jimmy Arias and Fernando Gonzales-Zubiate

Abstract

The use of synthetic biology on yeasts has enhanced the production of commercially relevant chemicals, from biofuels to recombinant therapeutic proteins, to name just a few. Despite most of these advances had already been studied and described in *Saccharomyces cerevisiae*, during the last years the attention has turned to the use of alternative expression systems with a higher yield and quality such as non-conventional yeasts. Recently, there has been an increase in studies about non-conventional yeasts due to advantages based on their natural capacity to tolerate harsh conditions or the wide range of carbon sources they need during the generation of specific products. This chapter, therefore, aims to describe the current status of the most used non-conventional yeasts in metabolite production as well as the engineering behind them in order to optimize or regulate protein expression: *Pichia pastoris*, *Kluyveromyces marxianus*, *Kluyveromyces lactis* and *Yarrowia lipolytica*.

Keywords: non-conventional yeasts, bioengineering, synthetic biology, CRISPR-Cas9, Golden Gate cloning, TALENs

1. Introduction

Yeast is probably one of the oldest domesticated organisms, since it was used for beer brewing already in Sumer and Babylonia around 6000 BC [1]. Not surprisingly, yeast cells were among the first microorganisms seen after the invention of the microscope in the seventeenth century, but their recognition as a living organisms did not come until two centuries later [2]. Yeasts, as such, do not form a single taxonomic or phylogenetic group in the kingdom fungi, rather, they occur in different subdivisions belonging to *Ascomycota and Basidiomycota*. Moreover, they are unicellular with budding and binary fission as the main asexual reproduction, and sexual spore production in stress conditions [3, 4].

Interestingly, yeast cells can exhibit a variety of cell sizes, shapes and colors. Even individual cells from a pure strain of a single species can display morphological heterogeneity. Moreover, yeast cell size varies widely from 2 up to 50 μm in length. Many yeast species are ellipsoidal or ovoid, but other cell shapes can be also observed as in *Candida albicans* and *Yarrowia lipolytica* which are mostly filamentous [1].

With respect to their diversity, there are around 2000 accepted yeast species included in the Yeast Trust Database (theyeasts.org) [4]. They have been isolated from highly diverse environments such as insect guts, food products, soil, oceans and even ancient ice fields [1]. However, it seems that we have just scratched the tip of the iceberg. According to Fell's estimation, what we have found represents only 1% of the species that might exist in nature [4].

Furthermore, yeasts as a whole are interesting because they are capable of metabolizing a wide variety of carbon sources including glucose, fructose, lactose, xylose and arabinose [3]. Besides, the metabolic activity of some yeasts can be dependent on the sugar concentration present in the medium: fermentation in high sugar concentration and aerobic respiration in low sugar concentrations (Crabtree effect) which can be advantageous in some industrial processes [3, 5].

2. Industrial applications of yeasts

Industrially, yeasts possess many attractive features that confer them some benefits in relation to bacteria such as *Escherichia coli*. For instance, yeasts have the capacity to grow on a wide variety of carbon sources, perform post-translational modifications, and compartmentalize reactions in organelles, they also present high secretion capacity, and are less susceptible to infectious agents like bacteriophages [6].

For these reasons, natural yeasts have been used in a lot of industrial processes. For example, in the food industry, the alcoholic fermentation of *S. cerevisiae* is used for the production of bread, beer and wine [7]. Furthermore, other yeasts species take different roles in the elaboration of food products such as yoghurts, in which *Torulopsis candida*, and *Kluyveromyces fragilis* are used for the improvement of aroma, texture and addition of nutrients by fermenting lactose with hydrolysis of milk casein [8].

Another important application of *S. cerevisiae* is the production of biofuels such as bioethanol, which is a result of sugar fermentation under anaerobic conditions. *S. cerevisiae* catabolizes sugars by glycolysis until it produces pyruvate that is then converted to acetaldehyde and carbon dioxide, which is reduced to ethanol by an alcohol dehydrogenase [9]. In addition, other yeast species also have the capacity to produce bioethanol. In fact, *Kluyveromyces marxianus*, *Dekkera bruxellensis* and *Scheffersomyces stipitis* are capable of producing bioethanol by fermentation of polyfructan substrates, hexoses and lignocelluloses substrates respectively [9]. Some yeasts are able to naturally produce bioethanol using lignocellulose resources (cheap, abundant and renewable) making them of great interest in second-generation biofuels, thus providing a clear advantage over first-generation biofuels that employ large cultivated areas [9–11].

On the other hand, despite *S. cerevisiae* has been widely studied and its industrial applications being countless, other yeast species, known as non-conventional yeasts (NCYs[1]) are becoming more popular in industrial applications. Several NCYs have diverse advantages compared to *S. cerevisiae*, mainly they are more suitable for a big number of biotechnological processes since they present natural tolerance to stresses like extreme pH, temperatures and osmolarity conditions [12]. Some of the most studied NCY species that are capable to withstand harsh conditions are *Yarrowia lipolytica*, *Hanensula polymorpha*, *Pichia pastoris* and *Kluyveromyces lactis* [13]. In fact, *K. lactis* is widely used in the cheese industry, replacing the conventional rennet, due

[1] Currently, there is not an accepted definition of NCYs, but many scientists consider NCYs as "non-*Saccharomyces*" yeasts [7].

to its ability to produce lactic acid from lactose. On the other hand, *Y. lipolytica* is used for the production of biosurfactants, carotenoids and lipids and *K. marxianus* for the production of bioethanol, aroma compounds and biosurfactants [6, 12].

Furthermore, some of the most interesting features of NCYs are their capacity to accumulate metabolites, synthesize and secrete recombinant proteins and enzymes [12]. For instance, yeasts like *Y. lipolytica* and *K. lactis* are able to secrete high titers of proteins extracellularly better than *S. cerevisiae* [13]. Moreover, some useful industrial enzymes like amylases, cellulases, proteases and lipases, have been reported to be produced by several strains of the NCY *Aureobasidium pullulans* [14]. These interesting properties have led to a proliferation of studies in NCYs aiming to improve their performance in the production of important metabolites and proteins. Hence, this review focuses on providing a clear description and analysis of the use of synthetic biology tools and strategies at the expression level that helps enhance four of the most popular NCYs: *P. pastoris*, *K. marxianus*, *K. lactis* and *Y. lipolytica* into valuable biofactories.

3. Use of synthetic biology in NCYs

Synthetic biology relies on the premise that a biological system can be built using a collection of previously described parts and subsystems [6]. This is achieved by standardization and modularization of useful biological parts, mechanisms and systems, or redesign the existing ones to provide new and better qualities [15]. Therefore, it is capable to define building blocks at various levels such as expression, protein and pathway levels [15]. The ability to control the dynamics at each level is important in order to establish unique and robust expression and production platforms for biomanufacturing [6]. For example, in synthetic biology-inspired therapies, the regulation of gene expression is important to determine the amount of the therapeutic and allows for accurate control over the design of synthetic cells [16].

Furthermore, with the emergence of modern genome editing tools, the synthetic biological capabilities to rewire and engineer organisms for production purposes have enabled the application of engineering efforts in non-conventional yeast of interest to industrial biotechnology [6].

This section will focus on the description of synthetic biology tools at the expression level, covering the engineering of genetic parts which include promoter, terminator and signal peptide, as well as codon optimization. In addition, available genome editing tools like CRISPR-Cas and cloning methods such as Golden Gate are also discussed. Their applications will be described in detail, later on, in Section 4.

3.1 Engineering of genetic parts

Codon optimization. The degeneracy of the genetic code means that several amino acids can be encoded by more than one codon (e.g. Leu = CUU, CUC, CUA, CUG), thus a random codon usage would be expected for those amino acids [17]. However, the expression of the same gene is different depending on the organism, because of the availability of host's tRNA pool. This is known as codon bias and is thought to affect the translation efficiency [18]. As a result, codon optimization is an important strategy when considering the expression of heterologous proteins.

Promoters. Selecting an adequate promoter is an important step since it can affect the level of expression of the desired protein. There can be constitutive or inducible

promoters, the latter more advantageous since they allow researchers to separate cell growth from the production of the desired protein. This avoids potential toxic effects due to a constitutive expression of heterologous proteins [19]. In addition, having a variety of promoters available is desired to fine-tune and optimize pathways that involve the de-expression of several proteins [20].

Terminators. It not only plays a critical role in transcription but is also able to influence mRNA stability and lifetime. This provides a new level of regulating protein expression; however, the impact of terminators is sometimes underestimated compared to promoters [21, 22]. In *S. cerevisiae* expression vectors, some endogenous terminators such as T_{CYC1} and T_{ADH1} are commonly used. Nevertheless, it has been demonstrated that *S. cerevisiae* terminators can show a high degree of transferability across other non-conventional yeasts [23].

Signal peptides. If a protein of interest is desired to be secreted, it is only required to add a secretion signal peptide at the N-terminus of the nascent polypeptide [24]. However, selecting an appropriate signal peptide is crucial since the protein quality and yield may vary widely depending on the heterologous protein being expressed [25]. Therefore, screening and characterizing many signal peptide sequences is a good approach to have adequate expression levels of different proteins.

3.2 Cloning methods and genome editing tools

Golden Gate. Golden Gate cloning method uses type II restriction enzymes (*Bsa*I and *Bpi*I) to precisely assemble multiple genetic parts by simultaneous restriction and ligation steps. These restriction enzymes cut outside their recognition sequences leaving 4-letter overhangs which can be freely designed; hence this method offers important benefits as it is cheaper than other advanced techniques, it does not require long flanking DNA and it allows scar-less cloning [26].

TALEN. Transcription-activator-like effector (TALE) nuclease is a genome editing tool based on type II effector proteins from bacterial plant pathogens of the genus *Xanthomonas* fused with the non-specific nuclease domain of the restriction endonuclease *Fok*I [27]. These TALENs are designed such that they bind separate targets in opposition to each other with an appropriate spacer between them allowing the *Fok*I nucleases to cause a double-strand break (DSB) and subsequently allowing the editing of the genome [28]. This genome editing tool has many advantages such as easy assembly, availability of powerful resources, cross-species flexibility and a high rate of success [27].

CRISPR-Cas9. This tool revolutionized the field of gene editing, since it is capable of creating a DSB in a specific DNA site by just using a single-guide RNA (sgRNA) complementary to the targeted region and an endonuclease (Cas9). This makes CRISPR-Cas9 a better genome editing tool compared to TALENs, which require laborious protein engineering steps for each new editing target. This editing tool has been successfully used for knock-out of a gene, but it can be adapted for other applications such as regulating transcription and facilitating metabolic engineering [29, 30].

4. Engineering non-conventional yeasts (NCYs)

This section will discuss the application of synthetic biology tools and strategies in four of the most popular NCYs: *P. pastoris*, *K. marxianus*, *K. lactis* and *Y. lipolytica*, yeasts that have been selected for several reasons. First, they naturally present

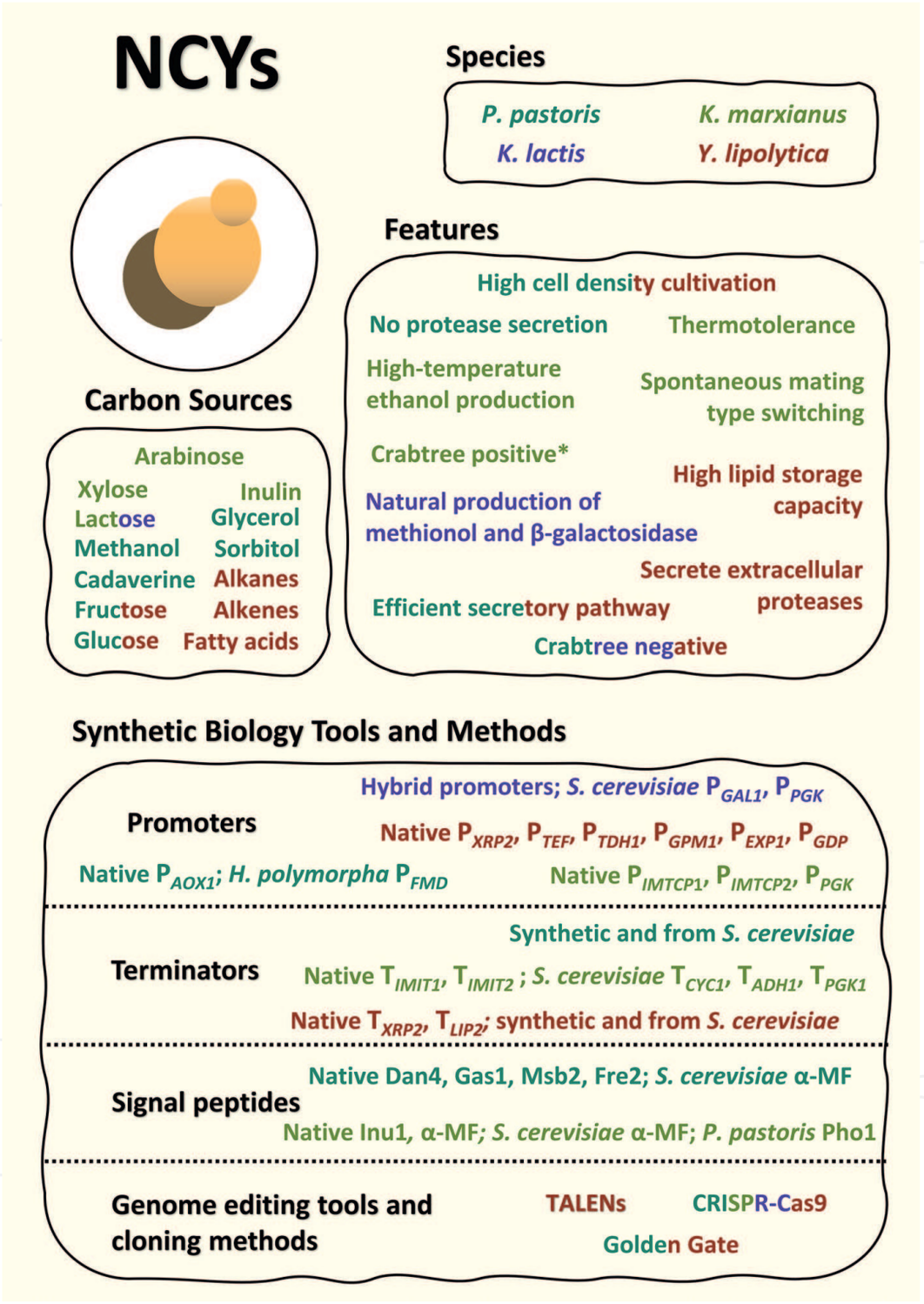

Figure 1.
Description of the non-conventional yeasts (NCYs): P. pastoris*,* K. marxianus*,* K. lactis *and* Y. lipolytica *and the synthetic biology tools and strategies applied to each of them. Descriptions of the same color represent a specific yeast. Example: All text in red belongs to* Y. lipolytica*.* *K. marxianus *is classified as Crabtree-negative, although some reports consider this species as Crabtree-positive due to strain variability.*

interesting characteristics with potential benefits in industrial applications such as the use of "waste products" as substrates, thermotolerance, high capacity to store lipids and the capacity to efficiently secrete desired compounds (**Figure 1**).

Second, they have their genomes sequenced and available at the National Center for Biotechnology Information (NCBI) website [31] and several metabolomic and transcriptomic studies have been reported [32–37]. This information is key when searching for new sources of BioBricks, since novel genetic parts or metabolites can be discovered under different contexts such as carbon sources or stress conditions. Third, information about regulatory associations between transcription factors (TFs) and target genes in these four species is available at the N. C. Yeastract database [38]. Last but not least, all the referred species have genome-scale metabolic models reported in the literature [39–42], allowing researchers to predict metabolic fluxes, and subsequently, optimize the production of relevant compounds in these microorganisms.

4.1 *Pichia pastoris*

Pichia pastoris[2] was initially developed by Phillips Petroleum Company for the production of single-cell protein for feedstock, but it was then repurposed as a promising expression system for the production of recombinant proteins [44] due to its many advantages such as high cell density cultivation, an efficient secretory capacity with a low background of endogenous proteins, the absence of protease secretion, the presence of alternative constitutive and inducible promoters and the ability to perform post-translational modifications to proteins [24, 45].

P. pastoris is a Crabtree-negative yeast that is able to utilize a variety of carbon sources including glucose, glycerol, fructose, sorbitol, methanol, alanine and cadaverine [43, 44]. As a matter of fact, it has the ability to use methanol as a sole carbon source (methylotrophic) due to several adaptations such as the expression of enzymes involved in methanol metabolism (e.g. alcohol oxidase) and the proliferation of peroxisomes (reaching over 80% of the cell volume) [44].

For these reasons, *P. pastoris* has become an industrially important microorganism. This is evidenced by the over 300 industrial processes that have been licensed and more than 70 commercial products that are currently on the market [46]. The use of synthetic biology tools in this yeast has led to a whole new level of potential industrial applications for this yeast. Some examples of the application of synthetic biology strategies in *P. pastoris* are provided below.

Codon Optimization. A clear application of this strategy in this yeast is the heterologous expression of enzymes such as phytases which are important enzymes that once included in animal feed can help not only increase the absorption of phosphorus in monogastric animals (e.g. pigs and horses) but at the same time reduce phosphorus levels in manure, thus finding a cheaper and more efficient way of producing this enzyme is desired. For instance, Xiong et al. compared the expression of a recombinant *Aspergillus niger* phytase in *P. pastoris* before and after codon optimization of both the phytase gene and the *S. cerevisiae* signal sequence α mating factor (α-MF). They obtained a phytase activity of 865 Units/ml, resulting in a 14.5-fold increase in the production/activity of phytase in comparison with the non-optimized gene and signal sequence [47]. The following year the same group applied this strategy to express a *Peniophora lycii* phytate enzyme achieving a phytase activity of 10,540 Units/ml and a 13.6-fold yield increase compared with the non-optimized gene and signal sequence [48]. In the last years, heterologous expression of other industrially

[2] Although initially named *Pichia pastoris* in the 1950s, it was then reclassified into the genus *Komagataella* in 1995, splitting up into the two species *K. pastoris* and *K. phaffii* [43]. Here, we still use the name *P. pastoris* for simplicity.

relevant proteins has been enhanced using this method, including but not limited to keratinases, endoinulinases, α-amylases, lipases, xylanases, fibases, pectinases, IFN-ω and hydroxynitrile lyases [49–57].

Promoters. Since all gene promoters of the methanol utilization (MUT) pathway are strongly repressed by carbon sources such as glucose, they can be a useful tool for the induced expression of heterologous proteins [58]. In fact, most of the heterologous expression of proteins in *P. pastoris* is carried out using one of those promoters, P*AOX*$_1$ [59]. In addition, other orthologous MUT promoters from related species have been evaluated in this species such as *Hansenula polymorpha* P_{FMD} which showed a 3.5-fold higher expression compared to the strongest endogenous MUT promoters [58].

On the other hand, using methanol (a flammable and toxic compound) in a large-scale fermentation process can be potentially dangerous [60]. For this reason, researchers have tried to find alternatives to overcome this limitation using different approaches such as employing orthologous promoters from related methylotrophic yeasts [58] or even engineering cis- and trans-acting elements in the P_{AOX1} [60]. For instance, using a trans-acting approach, Wang et al. developed a methanol-free method for protein expression using the P_{AOX1} promoter. They developed a strain that overexpressed the transcription activator Mit1 and repressed glucose- or glycerol-dependent transcriptional repressors Mig1, Mig2 and Nrg1. Then they evaluated the performance of both the methanol-free system and the wild-type system under their respective optimal culture conditions. Although the expression level of the recombinant insulin precursor protein in the methanol-free system was only 58.6% of the wild-type, they claim that it can be further improved by overexpressing unfolded protein response activators, protein foldases or chaperons [60].

Terminators. Knowledge about terminators is rather limited in *P. pastoris*, although in recent years it has been gaining more attention. In 2020 Ito et al. developed a terminator catalog of 72 sequences including synthetic, endogenous and heterogenous (*S. cerevisiae*) terminators with a 17-fold range of expression when using red fluorescent protein and *Aspergillus aculeatus* β-glucosidase as reporter proteins. Moreover, interesting results were found from this study including independence of terminator activity from the upstream gene and high degree of transferability of *S. cerevisiae* terminators to *P. pastoris* [61].

Signal Peptides. The *S. cerevisiae* signal sequence α-MF is widely used for secreted expression of recombinant proteins in *P. pastoris* [62], where a peptidic pre-region leads the translocation from the cytoplasm to the ER lumen. Additionally, it has a peptidic pro-region which facilitates the proper transit of target polypeptides from the ER to the Golgi apparatus; unfortunately, this pro-region tends to aggregate in the ER, impeding a proper secretion of the target protein [63]. Several strategies can be applied to overcome this issue such as codon optimization or directed evolution of the signal sequence α-MF [63]. Lately, new endogenous signal peptides could outperform signal sequence α-MF. For instance, in 2019 Duan and collaborators were able to identify four endogenous signal peptides (Dan4, Gas1, Msb2 and Fre2) in *P. pastoris*. All of them showed more than fourfold enhancement of total β-galactosidase activity over the signal sequence α-MF, with being Gas1 the one that showed the best results (230-fold increase). Moreover, Msb2 signal peptide had a better performance in the expression of β-galactosidase, yEGFP and cephalosporin C acylase; therefore, it could be considered as a more effective signal for heterologous protein secretion in *P. pastoris* [62].

Golden Gate. Prielhofer et al. developed GoldenPiCS a modular cloning system that facilitates the engineering of *P. pastoris* by generating episomal plasmids with up to 8 expression units. The main feature of this method is that researchers can easily

exchange genetic parts and quickly create and test new variants. The feasibility of this method was demonstrated with the optimization of a CRISPR-Cas9 system for *P. pastoris* using different combinations of humanized Cas9 and sgRNA on one single episomal plasmid [26].

In a more recent study, Cheng et al. used the Golden Gate cloning method to develop a versatile and easy way of assembling eukaryotic gene exons into both prokaryotic and eukaryotic plasmids in a one-step reaction [64]. Thus, this new approach enables researches to rapidly identify the optimal expression host for the production of specific proteins, overcoming some disadvantages of traditional methods to obtain intron-free eukaryotic genes (e.g. whole-gene synthesis or reverse transcription methods) which are time-consuming, expensive and complicated to operate [64].

CRISPR-Cas9. In *P. pastoris* exogenous cassettes with long homology arms are integrated ectopically and homologous recombination (HR) occurs only at variable frequencies of <0.1 to 30% [29]. However, high efficiency for gene insertion by HR can be achieved by deleting the protein Ku70 which is involved in nonhomologous end-joining (NHEJ) repair [65]. Weninger et al. used this strategy to develop integration cassettes of CRISPR-Cas9 marker-free with close to 100% efficiency [66]. Yang et al. developed a high-efficiency CRISPR-Cas system in *P. pastoris* synthetizing Cas9 (codon-optimized for *Homo sapiens)* and the sgRNA on different plasmids. They validated the editing efficiency in gene deletion of six genes, reaching or exceeding 75% of efficiency for each target gene. This system eliminates the episomal sgRNA plasmid through sub-culture to allow editing of another gene with a consistent single gene editing efficiency [65]. However, it was demonstrated that this system performed poorly when editing multiple genes. Finally, even though huge progress has been made to improve CRISPR-Cas9 efficiency in *P. pastoris*, it is still lower than other yeasts such as *S. cerevisiae* and *Kluyveromyces lactis* [65].

4.2 *Kluyveromyces marxianus*

K. marxianus is phylogenetically related to *S. cerevisiae* and more closely related to *K. lactis*. It is a hemiascomycetous yeast that can exist as stable haploid or diploid cells and is able to spontaneously switch its mating type (homothallic) [67, 68]. *K. marxianus* strains have been isolated from a great variety of habitats including dairy products, soil, sugarcane bagasse, insects and fruits. Therefore, this yeast presents a high metabolic diversity and a significant degree of intraspecific polymorphism [68–71], in point of fact, several industrially relevant compounds have been found to be naturally produced by *K. marxianus* including pectinase, aroma compounds, inulinase, lipase and lactase [69, 72]. Furthermore, some strains have been described to exhibit multi-stress resistance [68].

K. marxianus is classified as facultative fermentative and Crabtree negative, although some reports consider this species as Crabtree positive due to strain variability [69, 73]. *K. marxianus* cannot naturally grow under strictly anaerobic conditions [69], but it can be genetically modified to grow under such conditions. However, growth rates are still lower than anaerobic growth in *S. cerevisiae* [74]. *K. marxianus* is capable of using non-conventional sugars such as xylose, arabinose and inulin as carbon sources [40]. In addition, this yeast has the ability to use lactose which cannot be accomplished by *S. cerevisiae* and can grow at higher temperatures with a wider range of substrates than *K. lactis* [67, 69].

All these features make *K. marxianus* a promising biofactory, having a wide range of applications such as host for the production of heterologous proteins;

alternative to baker's yeast; bioremediation of textile dyes, cheese whey and copper; biomass for animal feeding; probiotics and high-temperature bioethanol production [67, 69, 72, 73, 75]. For instance, Nonklang et al. found remarkable differences in high-temperature ethanol production between *S. cerevisiae* and *K. marxianus* as *K. marxianus* DMKU3-1042 was the fastest to convert glucose to ethanol at 45°C whereas none of the *S. cerevisiae* strains were able to grow at this temperature [67].

In the last years, some progress in synthetic biology has been accomplished in this species; however, it is still limited compared to other NCYs such as *P. pastoris* since *K. marxianus* still lacks efficient genetic tools, there are also limited auxotrophic markers and very few constitutive and inducible promotes have been described [21]. Despite that, some examples and applications are discussed below.

Codon Optimization. This strategy is currently used to improve the expression of recombinant proteins in *K. marxianus*, especially vaccines. For instance, codon optimization has been used for the heterologous expression of the porcine circovirus type 2 (PCV2) Cap protein in *K. marxianus* as an alternative to produce PCV2 virus-like particle vaccines to treat porcine circovirus disease and help reduce economical losses in the swine industry. Duan et al. reported in their experiment higher yields compared with *E. coli* and *P. pastoris* as host vectors [76]. Other examples where codon optimization has been applied to *K. marxianus* include heterologous expression of single-chain antibodies, overproduction of inulinase, expression of the dengue virus type 1 nonstructural protein 1 and porcine parvovirus-like particles [77–80].

Promoters. In the last years, several promoters have been identified for this species which can be induced by heat, xylose, lactose or inulin [22]. In addition, several strong endogenous promoters of genes such as purine-cytosine permease, inulinase, enolase and glyceraldehyde 3-phosphate have been characterized [21, 69]. More recently, Kumar et al. identified two new strong promoters (P_{IMTCP1} and P_{IMTCP2}) which are more efficient at different temperatures and carbon sources than previously known promoters in this species [21].

Interestingly, the relative strength of promoters can change depending on the carbon source provided [81]. For instance, Kumar et al. showed that the *K. marxianus* inulinase promoter has relatively higher activity in the presence of xylose than dextrose [21]. In addition to finding new promoters, already described promoters have been also engineered to improve their features. For example, Zhou et al. improved the expression of lignocellulolytic enzymes in *K. marxianus* by a mutation inside the inulinase promoter and a deletion of an A-T-rich region inside the 5'UTR [82].

Notably, if thermotolerance of this yeast is to be exploited when expressing heterologous proteins, it is not enough to only focus on utilizing thermostable proteins but also identifying thermotolerant promoters because promoter activity tends to decrease with elevated temperature as it was demonstrated by Yang et al. [81]. Despite the fact that lower promoter activity is observed when increasing the temperature, *K. marxianus* promoters have been found to be stronger than their corresponding promoters from *S. cerevisiae* at such temperatures. In fact, the *K. marxianus* constitutive promoter P_{PGK} has been shown to retain relatively strong activity with an increase in temperature [81].

Terminators. Only a limited number of terminators have been examined in *K. marxianus* including terminators from *S. cerevisiae* such as T_{CYC1}, T_{PGK1} and T_{ADH1} [21, 22]. Additionally, new recently described *K. marxianus* terminators have widened the range of regulation of protein expression in this species. For instance, in a recent study researchers found an increase in EGFP expression (fourfold increase

of mRNA level) in *K. marxianus* when using the endogenous terminators T_{IMTT1} or T_{IMTT2} instead of *S. cerevisiae* T_{CYC1} [21].

Signal Peptides. Research about signal peptides in this species is still limited since only a few signal sequences have been characterized and employed in heterologous protein expression on *K. marxianus* including signal sequences from *K. marxianus* Inu1, *P. pastoris* Pho1, *S. cerevisiae* α-MF and *K. marxianus* α-MF [21]**.** Moreover, some experiments of signal sequence engineering conducted to improve its activity have been reported, for instance, Yarimizu et al. developed a synthetic signal sequence in the yeast *K. marxianus* by redesigning the hydrophobic core of *Gaussia princeps* secretory luciferase signal sequence. The hydrophobic sequence was replaced by a repeat of 16 methionine residues, resulting in 20-fold higher activity than that from the wild type [83].

CRISPR-Cas9. Cernak et al. first used a CRISPR-Cas9 system to inactivate genes responsible for spontaneous mating-type switching (common phenomenon in *K. marxianus*), enabling the production of stable heterothallic haploid strains which can mate. As a result, they combined three complex traits found in different strains (ability to take up exogenous DNA, thermotolerance capacity and higher lipid production) into single *K. marxianus* clones [71].

Li et al. developed a one-step multigene integration system based on CRISPR-Cas9, which is capable of integrating up to three cassettes in a single, targeted genomic locus in *K. marxianus.* It consists of the CRISPR plasmid (expression of the sgRNA and Cas9) and the homology donor plasmid (700 bp up- and down-stream homology to the targeted site). This system has been proven to have an efficiency comparable to single-gene integration and it can be performed within 4 days from transformation to confirmation [84].

In 2022 Bever et al. developed a highly efficient CRISPR-Cas9 system in *K. marxianus* that allows editing of multiple genes which can be used in both NHEJ-functional and -deficient strains showing nearly 100% efficiency of gene disruption in those two strains, whereas 100% efficiency of donor integration was observed only in NHEJ-deficient strains. In addition, this system achieved a dual integration efficiency of 25.5% in an NHEJ-deficient strain [85].

4.3 *Kluyveromyces lactis*

Kluyveromyces lactis is an NCY known for its capacity to assimilate lactose and convert it into lactic acid. *K. lactis* is a respiratory Crabtree-negative yeast highly used in industries due to its ability to secrete the protein β-galactosidase, used for making lactose-free products [23]. Moreover, this yeast is also capable to produce methionol, which is a flavor-active compound important in the overall aroma of soy sauce and cheese [7].

Some advantages of working with *K. lactis* yeasts are the capacity of producing heterologous proteins in simple growth medium, complete knowledge of their genome and more importantly, they can be easily genetically manipulated [75]. Due to its similarity in biosynthetic capacities to *S. cerevisiae*, *K. lactis* toolkits for heterologous gene expression are mostly the same. However, *K. lactis* presents many attributes that make it more suitable for protein expression and extracellular secretion than *S. cerevisiae* [13]. In fact, *K. lactis* uses an inducible promoter P_{LAC4} which is commercially available due to its capacity to secrete recombinant proteins in the culture fluid under lactose presence, a very useful feature for protein purification [13].

Promoters. In general, the promoters used for heterologous protein production strategies in *K. lactis* are the same as *S. cerevisiae*, P_{GAL1} or P_{PGK}, which, due to their

high level of transferability, have shown the potential of promoter engineering in *S. cerevisiae* to be applied in the *K. lactis* expression system [6]. However, other engineering strategies involving *K. lactis* promoters have also been developed. For example, Sakhtah et al. have recently developed a novel auto-inducible promoter system in *K. lactis*. For this, portions of two promoters, the constitutive P_{GAP1} and the carbon source-sensitive P_{ICL1}, were combined to form a hybrid promoter called P_{350} [86]. This novel promoter is induced by the depletion of glucose or glycerol in the medium, making it auto-inducible as the carbon sources are consumed by the growing cells. The development of this hybrid promoter promises to be useful for the implementation of one-step protein expression methods for small- and large-scale bioprocesses [86]. Moreover, another hybrid promoter approach used in K. *lactis* involves the combination of core promoter elements of *Trichoderma reesei* P_{CBH1} and *K. lactis* P_{LAC4}, which showed an increase in protein production in this yeast [13].

CRISPR-Cas9. For the implementation of CRISPR-Cas9 in *K. lactis*, Horwitz et al. adapted an *S. cerevisiae* system by exchanging the 2 μ element with the *K. lactis* specific pKD1 vector-stabilizing element and the constitutive promoter P_{FBA1} at a *GAL80* site [87–89]. Moreover, sgRNA expression was driven by the typical P_{SNR52} pol III promoter and a T*SUP*$_4$ terminator, and the deletion of the *KU80* gene was performed to reduce NHEJ [6, 29, 87–89]. The implementation of this system resulted in the successful integration of three donor six-gene-DNA parts into three separate loci (*DIT1, ADH1 and NDT80*) with a triple integration efficiency of 2.1% [6, 29, 87, 88]. Despite this low efficiency, the speed and ability to screen strains reduced the design-built-test cycle for this non-conventional yeast, and further improvements in targeting this efficiency could enhance genome editing for wild-type or industrial strains [6].

On the other hand, CRISPR-Cas9 genome editing was used by Burghardt et al. in order to increase the enzymatic production of the prebiotic fructo-oligosaccharides (FOS) in *K. lactis*. For this, the fructosyltransferase gene (*FFT)*, needed for forward reactions, from *Aspergillus terreus* NIH2624 was integrated with a *K. lactis* GG799 production host. Furthermore, a CRISPR-Cas9 system was used to delete a native invertase gene, involved in reverse reactions. The results showed an increase in transferase activity by 66.9% when grown in a fed-batch process [90].

4.4 *Yarrowia lipolytica*

Yarrowia lipolytica belongs to the *Ascomycota*, *Dipodascaceae* family. It is naturally found in lipid and protein-rich substrates such as soil, sewage and oil-polluted environments due to their capacity to hydrolyse lipids, assimilate hydrocarbons and fatty acids and secrete extracellular proteases [91]. *Y. lipolytica* is a Crabtree-negative haploid, heterothallic yeast with mating types Mat A and Mat B, and low mating frequency in nature [92]. In a laboratory setting, cells appear spherical, ellipsoidal or elongated and arranged singly, in pairs or clustered in groups. Furthermore, colonies present a creamy texture and a convoluted pale white matte surface [91]. About the carbon sources, *Y. lipolytica* is capable to assimilate hydrophobic substrates like alkanes, alkenes, fatty acids, fatty acid methyl esters, triglycerides and hydrophilic substrates like glucose, fructose, some alcohols, many polyols and many organic acids [91]. Some important characteristics of this yeast are its efficient secretion pathway and lipid storage capacity, two qualities that have made it a research model for protein secretion and lipid metabolism [92]. Moreover, due to its lipogenic metabolism, *Y. lipolytica* has been studied for the biosynthesis of acetyl-CoA-derived molecules such as terpenes [93]. In the last years, because of its production capacity of industrial

interest compounds and the ability to grow at high cell densities, different synthetic biology tools have been developed and applied in *Y. lipolytica* [94]. As matter of fact, Wong and colleagues designed a collection of BioBricks for *Y. lipolytica* called YaliBricks, which contains compatible restriction enzyme sites that allows modular genetic engineering [95].

Promoters. In *Yarrowia lipolytica*, two important promoters have been isolated and characterized; the promoter from the *XRP2* gene, which codes for an alkaline extracellular protease and the constitutive promoter from the *TEF* gene, which codes for translational elongation factor-1 [96]. Furthermore, recent studies are focusing on the development of hybrid promoters that could increase the strength of the available ones. Early approaches to promoter hybridization led to the characterization of upstream activating sequences (UASs), native to *XPR2*, which resulted in an increase on promoter activity when hybridized in several tandem repeats [97]. Madzak et al. engineered four hybrid promoters (named hp1d, hp2d, hp3d and hp4d) containing up to four copies of one of its upstream activation sequences ($UAS1_{XPR2}$) fused upstream from a P_{LEU2} promoter [98]. The resulting promoters showed an increase in their strength depending on the number of tandem $UAS1_{XPR2}$ elements, with hp4d being the strongest hybrid promoter and therefore used widely for heterologous gene expression in *Y. lipolytica* [94, 99]. On the other hand, Schwarts et al. constructed a synthetic hybrid promoter, using *GAL1* UAS from *S. cerevisiae* and the *TEF* core promoter from *Y. lipolytica*, that achieved a slightly higher expression than P*UAS1B8_TEF*, hybrid promoter that has been widely used [93]. Moreover, other native promoters like P_{TDH1}, P_{GPM1}, P_{EXP1}, P_{FBAINm}, P_{GPAT}, P_{GPD} and P_{YAT} have been characterized and used in expression of heterologous genes with promising results [94].

Terminators. The most commonly used terminators for expression of heterologous genes in *Y. lipolytica* are derived from the native *XPR2* and *LIP2* genes [99]. Moreover, some *S. cerevisiae* terminators have shown a high degree of transferability in *Y. lipolytica* [100]. Indeed, synthetic terminators designed for *S. cerevisiae* have been used in *Y. lipolytica* with an increase of 60% in expression level over some wild-type terminators [56]. Additionally, these synthetic terminators are commonly smaller than the natural ones, conferring them an advantage for transcription units (TU) and vector design since they show low risk of undesired HR between TU or with the genome, contributing high stability to genetically modified strains [50]. However, despite these advances, the number of studies of terminators in *Y. lipolytica*, in comparison to promoters, is still scarce [96].

Golden Gate. Larroude and collaborators have developed a modular toolkit based on the Golden Gate strategy that allows assembly in one step of three transcript units together with integration into *Y. lipolytica* genome. This approach comes with a collection of six selective markers and sequences for random or specific integration, nine promoters of variable strength and five terminators [100]. In such manner, the heterologous production of β-carotene is possible with the expression of three genes involved in the carotenoid pathway after a single transformation [101], making *Y. lipolytica* a competitive biotechnological producer of β-carotene.

CRISPR-Cas9. The use of CRISPR-Cas9 in *Y. lipolytica* has been widely studied. In recent years, a system from *Streptococcus pyogenes* has been adapted, with a synthetic RNA polymerase III promoter and an optimized Cas9 to perform a marker-free gene disruption and integration in *Y. lipolytica*. In fact, five loci have been recently identified that could serve as hotspots for targeting marker-free gene integration [97]. This system resulted in a single-gene disruption and HR with a 90 and 70% of efficiency

respectively when Cas9 and the sgRNA were co-transformed using donor DNA [94]. In addition, systems like CRISPR interference (CRISPRi) and CRISPR activation (CRISPRa) have been developed for controlling gene expression in *Y. lipolytica* [99]. Here, a deactivated Cas9 (dCas9) is fused to transcriptional repressors or activators, allowing binding to sgRNA-complementary DNA without cleavage that could result in DSBs [91]. An implementation of the CRISPRa system was performed by Schwarts et al. in order to upregulate *BGL1* and *BGL2* (β-glucosidase genes that are transcriptionally silent), and allow *Y. lipolytica* strains to use cellobiose as a carbon source. For this, a VPR activator was identified and fused to dCas9 to enable gene activation [93, 96].

Other CRISPR-Cas9 strategies have been developed, for example, a paired sgRNA, consisting of two vectors, each containing Cas9 gene and a sgRNA cassette, was used in order to target areas upstream and downstream the start and stop codon, respectively, and allow a complete gene knockout via gene excision with a 20% of efficiency [94].

TALEN. Used to direct DNA DSBs to occur at a specific target site, was applied in *Y. lipolytica* to produce structure-based mutagenesis of a fatty acid synthase (FAS) domain, and allow the synthesis of fatty acids with shorter chain lengths [99]. Moreover, site-directed mutagenesis improved in efficiency when homologous exogenous DNA was added to the targeted site, resulting in HR-mediated repair in 40% of clones [94].

5. Future perspectives

To date, only about 1% of the yeast species found in nature have had their genomes fully characterized. Thus, it is not surprising, with the accessibility of new sequencing technologies, the complete genome analysis of many newly discovered yeast species with unique characteristics will be available. This will greatly expand the catalog of genetic parts allowing a more sensitive fine-tuning of desired economically relevant compounds and the discovery of new genes of interest. In addition, new yeast host vectors with desirable characteristics such as faster growth rates, stress-tolerance, efficient secretion systems and desired metabolic pathways will be engineered and domesticated to facilitate their use in industrial applications.

Moreover, in silico simulations will play a crucial role in the efficient design of new synthetic yeast biofactories since more accurate predictions will be made. Nevertheless, there are still limitations that have to be overcome such as the absence of gene regulatory information, lack of accurate metabolic models at genomic scale [102], or missing experimental design and testing of potential NCY biofactories. For instance, CRISPR-Cas9 still performs poorly regarding the adequate sgRNA production in NCYs. sgRNA expression is normally accomplished using RNA polymerase III (RNAP III) promoters (not well characterized in NCYs), implying more studies are needed for effective genetic engineering. Wagner and Alper suggest two approaches to overcome this issue: optimization of heterologous RNAP III promoters or the screening of native RNAP III promoters [6, 13] which are currently being tested in some NCYs such as *Y. lipolytica* [97].

Finally, a quite positive outcome of the use of NCYs as biofactories is the production of industrially relevant compounds in an economical manner. Synthetic biology helps to search and engineer strains capable of utilizing cheaper substrates, including "waste products" (e.g. whey and molasses), supporting a sustainable circular economy which in the future will certainly have a more relevant role.

6. Conclusions

For several decades yeasts have proven to be of great importance for the development of modern society, contributing to industrial processes including food and pharmaceutics. In addition, the current application of synthetic biology techniques in these organisms has given them a greater potential to be used as substitutes for organisms commonly used in the industry, such as bacteria, given the benefits they present. The use of these techniques in unconventional yeasts such as *P. pastoris*, *K. marxianus*, *K. lactis* and *Y. lipolytica* has increased very rapidly in recent years. For example, the genome editing technique CRISPR-Cas9 has been developed in these four species to improve the production of compounds, such as the prebiotic fructooligosaccharides in the case of *K. lactis*. Other techniques, despite being recently applied, have shown promising results in improving the expression of genes and the production of compounds of interest. This is the case of the use of TALENs in *Y. lipolytica* and Golden Gate in *P. pastoris*.

On the other hand, the engineering of genetic parts has also been developed in these unconventional yeasts. Codon optimization in *P. pastoris* and *K. marxianus* has allowed the production of heterologous enzymes such as phytases, important in animal feed, and recombinant proteins such as PCV2 Cap protein, which can be used in vaccines against porcine circovirus disease. Modifications of promoters and terminators have also been investigated, with promoter studies being the most common, as shown by the recent literature existing on the four species described in this chapter. In sum, the advances shown here demonstrate the potential of non-conventional yeasts as alternatives to traditionally used organisms, or even for the discovery of new systems with potential industrial use, capable of improving the quality of people's life.

Author details

Kevin Castillo-Mendieta[1], Jimmy Arias[1] and Fernando Gonzales-Zubiate[1,2*]

1 School of Biological Sciences and Engineering, Yachay Tech University, Urcuquí, Ecuador

2 MIND Research Group, Model Intelligent Networks Development, Urcuquí, Ecuador

*Address all correspondence to: fgonzales@yachaytech.edu.ec

References

[1] Feldmann H. Yeast Molecular and Cell Biology. Weinheim, Germany: Wiley-Blackwell; 2012

[2] Nicholas P. Money. The Rise of Yeast: Oxford University Press; 2018

[3] Satyanarayana T, Kunze G. Yeast Diversity in Human Welfare. Singapore: Springer; 2017. DOI: 10.1007/978-981-10-2621-8

[4] Boekhout T, Amend AS, el Baidouri F, Gabaldón T, Geml J, Mittelbach M, et al. Trends in yeast diversity discovery. Fungal Diversity. 2021;**114**:491-537. DOI: 10.1007/S13225-021-00494-6

[5] de Deken RH. The Crabtree effect: A regulatory system in yeast. Journal of General Microbiology. 1966;**44**:149-156. DOI: 10.1099/00221287-44-2-149

[6] Wagner JM, Alper HS. Synthetic biology and molecular genetics in non-conventional yeasts: Current tools and future advances. Fungal Genetics and Biology. 2016;**89**:126-136. DOI: 10.1016/j.fgb.2015.12.001

[7] Forti L, Cramarossa MR, Filippucci S, Tasselli G, Turchetti B, Buzzini P. Nonconventional yeast-promoted biotransformation for the production of flavor compounds. Handbook of Food Bioengineering. 2018;**7**:165-187. DOI: 10.1016/B978-0-12-811518-3.00006-5

[8] Nandy SK, Srivastava RK. A review on sustainable yeast biotechnological processes and applications. Microbiological Research. 2018;**207**:83-90. DOI: 10.1016/j.micres.2017.11.013

[9] Parapouli M, Vasileiadis A, Afendra AS, Hatziloukas E. *Saccharomyces cerevisiae* and its industrial applications. AIMS Microbiology. 2020;**6**:1-31. DOI: 10.3934/microbiol.2020001

[10] Unrean P, Khajeeram S. Model-based optimization of *Scheffersomyces stipitis* and *Saccharomyces cerevisiae* co-culture for efficient lignocellulosic ethanol production. Bioresources and Bioprocessing. 2015;**2**:1-11. DOI: 10.1186/S40643-015-0069-1

[11] Lignocellulose SK. A chewy problem. Nature. 2011;**474**:S12-S14. DOI: 10.1038/474s012a

[12] Geijer C, Ledesma-Amaro R, Tomas-Pejo E. Unraveling the potential of non-conventional yeasts in biotechnology. FEMS Yeast Research. 2022;**22**:1-6. DOI: 10.1093/femsyr/foab071

[13] Madhavan A, Jose AA, Binod P, Sindhu R, Sukumaran RK, Pandey A, et al. Synthetic biology and metabolic engineering approaches and its impact on non-conventional yeast and biofuel production. Frontiers in Energy Research. 2017;**5**:1-12. DOI: 10.3389/fenrg.2017.00008

[14] Prasongsuk S, Lotrakul P, Ali I, Bankeeree W, Punnapayak H. The current status of Aureobasidium pullulans in biotechnology. Folia Microbiologia (Praha). 2018;**63**:129-140. DOI: 10.1007/s12223-017-0561-4

[15] Liu Z, Zhang Y, Nielsen J. Synthetic biology of yeast. Biochemistry. 2019;**58**:1511-1520. DOI: 10.1021/acs.biochem.8b01236

[16] Bojar D, Scheller L, El HGC, Xie M, Fussenegger M. Caffeine-inducible gene switches controlling experimental

diabetes. Nature Communications. 2018;**9**:1-10. DOI: 10.1038/s41467-018-04744-1

[17] Kurland CG. Codon bias and gene expression. FEBS Letters. 1991;**285**:165-169. DOI: 10.1016/0014-5793(91)80797-7

[18] Quax TEF, Claassens NJ, Söll D, van der Oost J. Codon bias as a means to fine-tune gene expression. Molecular Cell. 2015;**59**:149-161. DOI: 10.1016/J.MOLCEL.2015.05.035

[19] Türkanoğlu Özçelik A, Yılmaz S, Inan M. *Pichia pastoris* promoters. Methods in Molecular Biology. 2019;**1923**:97-112. DOI: 10.1007/978-1-4939-9024-5_3

[20] Vogl T, Sturmberger L, Kickenweiz T, Wasmayer R, Schmid C, Hatzl AM, et al. A toolbox of diverse promoters related to methanol utilization: Functionally verified parts for heterologous pathway expression in *Pichia pastoris*. ACS Synthetic Biology. 2016;**5**:172-186. DOI: 10.1021/acssynbio.5b00199

[21] Kumar P, Sahoo DK, Sharma D. The identification of novel promoters and terminators for protein expression and metabolic engineering applications in *Kluyveromyces marxianus*. Metabolic Engineering Communications. 2021;**12**:1-13. DOI: 10.1016/J.MEC.2020.E00160

[22] Rajkumar AS, Varela JA, Juergens H, Daran JMG, Morrissey JP. Biological parts for *Kluyveromyces marxianus* synthetic biology. Frontiers in Bioengineering and Biotechnology. 2019;**7**:97. DOI: 10.3389/fbioe.2019.00097

[23] Gomes AMV, Carmo TS, Carvalho LS, Bahia FM, Parachin NS. Comparison of yeasts as hosts for recombinant protein production. Microorganisms. 2018;**6**:38. DOI: 10.3390/microorganisms6020038

[24] Obst U, Lu TK, Sieber V. A modular toolkit for generating *Pichia pastoris* secretion libraries. ACS Synthetic Biology. 2017;**6**:1016-1025. DOI: 10.1021/acssynbio.6b00337

[25] Dou W, Zhu Q, Zhang M, Jia Z, Guan W. Screening and evaluation of the strong endogenous promoters in *Pichia pastoris*. Microbial Cell Factories. 2021;**20**:1-12. DOI: 10.1186/S12934-021-01648-6

[26] Prielhofer R, Barrero JJ, Steuer S, Gassler T, Zahrl R, Baumann K, et al. GoldenPiCS: A golden gate-derived modular cloning system for applied synthetic biology in the yeast *Pichia pastoris*. BMC Systems Biology. 2017;**11**:1-14. DOI: 10.1186/S12918-017-0492-3

[27] Wright DA, Li T, Yang B, Spalding MH. TALEN-mediated genome editing: Prospects and perspectives. Biochemical Journal. 2014;**462**:15-24. DOI: 10.1042/BJ20140295

[28] Li T, Wright DA, Spalding MH, Yang B. TALEN-based Genome Editing in Yeast. Cham, Switzerland: Springer 2015. pp. 289-307. DOI: 10.1007/978-3-319-10142-2_27

[29] Raschmanová H, Weninger A, Glieder A, Kovar K, Vogl T. Implementing CRISPR-Cas technologies in conventional and non-conventional yeasts: Current state and future prospects. Biotechnology Advances. 2018;**36**:641-665. DOI: 10.1016/J.BIOTECHADV.2018.01.006

[30] Yip BH. Recent advances in CRISPR/Cas9 delivery strategies. Biomolecules. 2020;**10**:839. DOI: 10.3390/BIOM10060839

[31] Sayers EW, Bolton EE, Brister JR, Canese K, Chan J, Comeau DC, et al. Database resources of the national center

for biotechnology information. Nucleic Acids Research. 2022;**50**:D20-D26. DOI: 10.1093/NAR/GKAB1112

[32] Love KR, Shah KA, Whittaker CA, Wu J, Bartlett MC, Ma D, et al. Comparativegenomicsandtranscriptomics of *Pichia pastoris*. BMC Genomics. 2016;**17**:1-17. DOI: 10.1186/S12864-016-2876-Y

[33] Nurcholis M, Lertwattanasakul N, Rodrussamee N, Kosaka T, Murata M, Yamada M. Integration of comprehensive data and biotechnological tools for industrial applications of *Kluyveromyces marxianus*. Applied Microbiology and Biotechnology. 2020;**104**:475-488. DOI: 10.1007/S00253-019-10224-3

[34] Becerra M, González-Siso MI, Cerdán ME. A transcriptome analysis of *Kluyveromyces lactis* growing in cheese whey. International Dairy Journal. 2006;**16**:207-214. DOI: 10.1016/J.IDAIRYJ.2005.03.005

[35] Morin N, Cescut J, Beopoulos A, Lelandais G, le Berre V, Uribelarrea JL, et al. Transcriptomic analyses during the transition from biomass production to lipid accumulation in the oleaginous yeast yarrowia lipolytica. PLoS One. 2011;**6**:e27966. DOI: 10.1371/JOURNAL.PONE.0027966

[36] Pomraning KR, Wei S, Karagiosis SA, Kim YM, Dohnalkova AC, Arey BW, et al. Comprehensive metabolomic, lipidomic and microscopic profiling of yarrowia lipolytica during lipid accumulation identifies targets for increased lipogenesis. PLoS One. 2015;**10**:1-17. DOI: 10.1371/JOURNAL.PONE.0123188

[37] Alvim MCT, Vital CE, Barros E, Vieira NM, da Silveira FA, Balbino TR, et al. Ethanol stress responses of *Kluyveromyces marxianus* CCT 7735 revealed by proteomic and metabolomic analyses. Antonie van Leeuwenhoek, International Journal of General and Molecular Microbiology. 2019;**112**:827-845. DOI: 10.1007/S10482-018-01214-Y

[38] Godinho CP, Palma M, Oliveira J, Mota MN, Antunes M, Teixeira MC, et al. The N.C.yeastract and communityyeastract databases to study gene and genomic transcription regulation in non-conventional yeasts. FEMS Yeast Research. 2021;**21**:1-10. DOI: 10.1093/FEMSYR/FOAB045

[39] Ye R, Huang M, Lu H, Qian J, Lin W, Chu J, et al. Comprehensive reconstruction and evaluation of *Pichia pastoris* genome-scale metabolic model that accounts for 1243 ORFs. Bioresour Bioprocess. 2017;**4**:1-12. DOI: 10.1186/S40643-017-0152-X

[40] Marcišauskas S, Ji B, Nielsen J. Reconstruction and analysis of a *Kluyveromyces marxianus* genome-scale metabolic model. BMC Bioinformatics. 2019;**20**:1-9. DOI: 10.1186/S12859-019-3134-5

[41] Dias O, Pereira R, Gombert AK, Ferreira EC, Rocha I. iOD907, the first genome-scale metabolic model for the milk yeast *Kluyveromyces lactis*. Biotechnology Journal. 2014;**9**:776-790. DOI: 10.1002/BIOT.201300242

[42] Xu Y, Holic R, Hua Q. Comparison and analysis of published genome-scale metabolic models of yarrowia lipolytica. Biotechnology and Bioprocess Engineering. 2020;**25**:53-61. DOI: 10.1007/S12257-019-0208-1

[43] Peña DA, Gasser B, Zanghellini J, Steiger MG, Mattanovich D. Metabolic engineering of *Pichia pastoris*. Metabolic Engineering. 2018;**50**:2-15. DOI: 10.1016/J.YMBEN.2018.04.017

[44] Sreekrishna K, Kropp KE. *Pichia pastoris*. In: Nonconventional Yeasts in Biotechnology. 1996. pp. 203-253. DOI: 10.1007/978-3-642-79856-6_6

[45] Kang Z, Huang H, Zhang Y, Du G, Chen J. Recent advances of molecular toolbox construction expand *Pichia pastoris* in synthetic biology applications. World Journal of Microbiology and Biotechnology. 2016;**33**:1-8. DOI: 10.1007/S11274-016-2185-2

[46] García-Ortega X, Cámara E, Ferrer P, Albiol J, Montesinos-Seguí JL, Valero F. Rational development of bioprocess engineering strategies for recombinant protein production in *Pichia pastoris* (Komagataella phaffii) using the methanol-free GAP promoter. Where do we stand? New Biotechnology. 2019;**53**:24-34. DOI: 10.1016/J.NBT.2019.06.002

[47] Xiong AS, Yao QH, Peng RH, Han PL, Cheng ZM, Li Y. High level expression of a recombinant acid phytase gene in *Pichia pastoris*. Journal of Applied Microbiology. 2005;**98**:418-428. DOI: 10.1111/J.1365-2672.2004.02476.X

[48] Xiong AS, Yao QH, Peng RH, Zhang Z, Xu F, Liu JG, et al. High level expression of a synthetic gene encoding Peniophora lycii phytase in methylotrophic yeast *Pichia pastoris*. Applied Microbiology and Biotechnology. 2006;**72**:1039-1047. DOI: 10.1007/S00253-006-0384-8

[49] Hu H, Gao J, He J, Yu B, Zheng P, Huang Z, et al. Codon optimization significantly improves the expression level of a keratinase gene in *Pichia pastoris*. PLoS One. 2013;**8**:1-8. DOI: 10.1371/JOURNAL.PONE.0058393

[50] Wang Y, Jiang S, Jiang X, Sun X, Guan X, Han Y, et al. Cloning and codon optimization of a novel feline interferon omega gene for production by *Pichia pastoris* and its antiviral efficacy in polyethylene glycol-modified form. Virulence. 2022;**13**:297-309. DOI: 10.1080/21505594.2022.2029330

[51] He M, Wu D, Wu J, Chen J. Enhanced expression of endoinulinase from Aspergillus niger by codon optimization in *Pichia pastoris* and its application in inulooligosaccharide production. Journal of Industrial Microbiology & Biotechnology. 2014;**41**:105-114. DOI: 10.1007/S10295-013-1341-Z

[52] Wang JR, Li YY, Liu DN, Liu JS, Li P, Chen LZ, et al. Codon optimization significantly improves the expression level of α-Amylase gene from Bacillus licheniformis in *Pichia pastoris*. BioMed Research International. 2015;**2015**:1-9. DOI: 10.1155/2015/248680

[53] Ahn J, Jang MJ, Ang KS, Lee H, Choi ES, Lee DY. Codon optimization of Saccharomyces cerevisiae mating factor alpha prepro-leader to improve recombinant protein production in *Pichia pastoris*. Biotechnology Letters. 2016;**38**:2137-2143. DOI: 10.1007/S10529-016-2203-3

[54] Hu H, Dai S, Wen A, Bai X. Efficient expression of xylanase by codon optimization and its effects on the growth performance and carcass characteristics of broiler. Animals. 2019;**9**:65. DOI: 10.3390/ANI9020065

[55] Che Z, Cao X, Chen G, Liang Z. An effective combination of codon optimization, gene dosage, and process optimization for high-level production of fibrinolytic enzyme in Komagataella phaffii (*Pichia pastoris*). BMC Biotechnology. 2020;**20**:1-13. DOI: 10.1186/S12896-020-00654-7

[56] Karaoğlan M, Erden-Karaoğlan F. Effect of codon optimization and

promoter choice on recombinant endo-polygalacturonase production in *Pichia pastoris*. Enzyme and Microbial Technology. 2020;**139**:1-9. DOI: 10.1016/J.ENZMICTEC.2020.109589

[57] Zhai Z, Nuylert A, Isobe K, Asano Y. Effects of codon optimization and glycosylation on the high-level production of hydroxynitrile lyase from Chamberlinius hualienensis in *Pichia pastoris*. Journal of Industrial Microbiology & Biotechnology. 2019;**46**:887-898. DOI: 10.1007/S10295-019-02162-W

[58] Vogl T, Fischer JE, Hyden P, Wasmayer R, Sturmberger L, Glieder A. Orthologous promoters from related methylotrophic yeasts surpass expression of endogenous promoters of *Pichia pastoris*. AMB Express. 2020;**10**:1-9. DOI: 10.1186/S13568-020-00972-1

[59] Nong L, Zhang Y, Duan Y, Hu S, Lin Y, Liang S. Engineering the regulatory site of the catalase promoter for improved heterologous protein production in *Pichia pastoris*. Biotechnology Letters. 2020;**42**:2703-2709. DOI: 10.1007/S10529-020-02979-X

[60] Wang J, Wang X, Shi L, Qi F, Zhang P, Zhang Y, et al. Methanol-independent protein expression by AOX1 promoter with trans-acting elements engineering and glucose-glycerol-shift induction in *Pichia pastoris*. Scientific Reports. 2017;**7**:1-12. DOI: 10.1038/srep41850

[61] Ito Y, Terai G, Ishigami M, Hashiba N, Nakamura Y, Bamba T, et al. Exchange of endogenous and heterogeneous yeast terminators in *Pichia pastoris* to tune mRNA stability and gene expression. Nucleic Acids Research. 2020;**48**:13000-13012. DOI: 10.1093/NAR/GKAA1066

[62] Duan G, Ding L, Wei D, Zhou H, Chu J, Zhang S, et al. Screening endogenous signal peptides and protein folding factors to promote the secretory expression of heterologous proteins in *Pichia pastoris*. Journal of Biotechnology. 2019;**306**:193-202. DOI: 10.1016/J.JBIOTEC.2019.06.297

[63] Ito Y, Ishigami M, Hashiba N, Nakamura Y, Terai G, Hasunuma T, et al. Avoiding entry into intracellular protein degradation pathways by signal mutations increases protein secretion in *Pichia pastoris*. Microbial Biotechnology. 2022;**15**:2364-2378. DOI: 10.1111/1751-7915.14061

[64] Cheng J, Wu M, Zhong R, Si D, Meng G, Zhang R, et al. Rapid golden gate assembly of exons from genomic DNA for protein expression in Escherichia coli and *Pichia pastoris*. BioTechniques. 2021;**71**:445-450. DOI: 10.2144/BTN-2021-0039

[65] Yang Y, Liu G, Chen X, Liu M, Zhan C, Liu X, et al. High efficiency CRISPR/Cas9 genome editing system with an eliminable episomal sgRNA plasmid in *Pichia pastoris*. Enzyme and Microbial Technology. 2020;**138**:109556. DOI: 10.1016/J.ENZMICTEC.2020.109556

[66] Weninger A, Fischer JE, Raschmanová H, Kniely C, Vogl T, Glieder A. Expanding the CRISPR/Cas9 toolkit for *Pichia pastoris* with efficient donor integration and alternative resistance markers. Journal of Cellular Biochemistry. 2018;**119**:3183-3198. DOI: 10.1002/JCB.26474

[67] Nonklang S, Abdel-Banat BMA, Cha-aim K, Moonjai N, Hoshida H, Limtong S, et al. High-temperature ethanol fermentation and transformation with linear DNA in the thermotolerant yeast *Kluyveromyces marxianus* DMKU3-1042. Applied and Environmental Microbiology. 2008;**74**:7514-7521. DOI: 10.1128/AEM.01854-08

[68] Lane MM, Burke N, Karreman R, Wolfe KH, O'Byrne CP, Morrissey JP. Physiological and metabolic diversity in the yeast *Kluyveromyces marxianus*. Antonie Van Leeuwenhoek. 2011;**100**:507-519. DOI: 10.1007/S10482-011-9606-X

[69] Fonseca GG, Heinzle E, Wittmann C, Gombert AK. The yeast *Kluyveromyces marxianus* and its biotechnological potential. Applied Microbiology and Biotechnology. 2008;**79**:339-354. DOI: 10.1007/S00253-008-1458-6

[70] Morrissey JP, Etschmann MMW, Schrader J, de Billerbeck GM. Cell factory applications of the yeast *Kluyveromyces marxianus* for the biotechnological production of natural flavour and fragrance molecules. Yeast. 2015;**32**:3-16. DOI: 10.1002/YEA.3054

[71] Cernak P, Estrela R, Poddar S, Skerker JM, Cheng YF, Carlson AK, et al. Engineering *Kluyveromyces marxianus* as a robust synthetic biology platform host. MBio. 2018;**9**:1-16. DOI: 10.1128/MBIO.01410-18

[72] Karim A, Gerliani N, Aïder M. *Kluyveromyces marxianus*: An emerging yeast cell factory for applications in food and biotechnology. International Journal of Food Microbiology. 2020;**333**:108818. DOI: 10.1016/J.IJFOODMICRO.2020.108818

[73] Lane MM, Morrissey JP. *Kluyveromyces marxianus*: A yeast emerging from its sister's shadow. Fungal Biology Reviews. 2010;**24**:17-26. DOI: 10.1016/J.FBR.2010.01.001

[74] Dekker WJC, Ortiz-Merino RA, Kaljouw A, Battjes J, Wiering FW, Mooiman C, et al. Engineering the thermotolerant industrial yeast *Kluyveromyces marxianus* for anaerobic growth. Metabolic Engineering. 2021;**67**:347-364. DOI: 10.1016/J.YMBEN.2021.07.006

[75] Bilal M, Ji L, Xu Y, Xu S, Lin Y, Iqbal HMN, et al. Bioprospecting *Kluyveromyces marxianus* as a Robust Host for Industrial Biotechnology. Frontiers in Bioengineering and Biotechnology. 2022;**10**:562. DOI: 10.3389/FBIOE.2022.851768

[76] Duan J, Yang D, Chen L, Yu Y, Zhou J, Lu H. Efficient production of porcine circovirus virus-like particles using the nonconventional yeast *Kluyveromyces marxianus*. Applied Microbiology and Biotechnology. 2018;**103**:833-842. DOI: 10.1007/S00253-018-9487-2

[77] Yang D, Chen L, Duan J, Yu Y, Zhou J, Lu H. Investigation of *Kluyveromyces marxianus* as a novel host for large-scale production of porcine parvovirus virus-like particles. Microbial Cell Factories. 2021;**20**:1-13. DOI: 10.1186/S12934-021-01514-5

[78] Bragança CRS, Colombo LT, Roberti AS, Alvim MCT, Cardoso SA, Reis KCP, et al. Construction of recombinant *Kluyveromyces marxianus* UFV-3 to express dengue virus type 1 nonstructural protein 1 (NS1). Applied Microbiology and Biotechnology. 2015;**99**:1191-1203. DOI: 10.1007/S00253-014-5963-5

[79] Nambu-Nishida Y, Nishida K, Hasunuma T, Kondo A. Genetic and physiological basis for antibody production by *Kluyveromyces marxianus*. AMB Express. 2018;**8**:1-9. DOI: 10.1186/S13568-018-0588-1

[80] Zhang Y, Li YF, Chi Z, Liu GL, Jiang H, Hu Z, et al. Inulinase hyperproduction by *Kluyveromyces marxianus* through codon optimization, selection of the promoter, and high-cell-density fermentation for efficient inulin

hydrolysis. Annales de Microbiologie. 2019;**69**:647-657. DOI: 10.1007/S13213-019-01457-8

[81] Yang C, Hu S, Zhu S, Wang D, Gao X, Hong J. Characterizing yeast promoters used in *Kluyveromyces marxianus*. World Journal of Microbiology and Biotechnology. 2015;**31**:1641-1646. DOI: 10.1007/S11274-015-1899-X

[82] Zhou J, Zhu P, Hu X, Lu H, Yu Y. Improved secretory expression of lignocellulolytic enzymes in *Kluyveromyces marxianus* by promoter and signal sequence engineering. Biotechnology for Biofuels. 2018;**11**:1-14. DOI: 10.1186/S13068-018-1232-7

[83] Yarimizu T, Nakamura M, Hoshida H, Akada R. Synthetic signal sequences that enable efficient secretory protein production in the yeast *Kluyveromyces marxianus*. Microbial Cell Factories. 2015;**14**:1-14. DOI: 10.1186/S12934-015-0203-Y

[84] Li M, Lang X, Moran Cabrera M, de Keyser S, Sun X, da Silva N, et al. CRISPR-mediated multigene integration enables Shikimate pathway refactoring for enhanced 2-phenylethanol biosynthesis in *Kluyveromyces marxianus*. Biotechnology for Biofuels. 2021;**14**:1-15. DOI: 10.1186/S13068-020-01852-3

[85] Bever D, Wheeldon I, da Silva N. RNA polymerase II-driven CRISPR-Cas9 system for efficient non-growth-biased metabolic engineering of *Kluyveromyces marxianus*. Metabolic Engineering Communications. 2022:1-10. DOI: 10.1016/J.MEC.2022.E00208

[86] Sakhtah H, Behler J, Ali-Reynolds A, Causey TB, Vainauskas S, Taron CH. A novel regulated hybrid promoter that permits autoinduction of heterologous protein expression in *Kluyveromyces lactis*. Applied and Environmental Microbiology. 2019;**85**:1-12. DOI: 10.1128/AEM.00542-19

[87] Horwitz AA, Walter JM, Schubert MG, Kung SH, Hawkins K, Platt DM, et al. Efficient multiplexed integration of synergistic alleles and metabolic pathways in yeasts via CRISPR-Cas. Cell Systems. 2015;**1**:88-96. DOI: 10.1016/j.cels.2015.02.001

[88] Cai P, Gao J, Zhou Y. CRISPR-mediated genome editing in non-conventional yeasts for biotechnological applications. Microbial Cell Factories. 2019;**18**:1-12. DOI: 10.1186/s12934-019-1112-2

[89] Bisschoff E. The Development of a Wide Range CRISPR-Cas9 Gene Editing System. Bloemfontein, South Africa: University of the Free State; 2019

[90] Burghardt JP, Fan R, Baas M, Eckhardt D, Gerlach D, Czermak P. Enhancing the heterologous fructosyltransferase activity of *Kluyveromyces lactis*: Developing a scaled-up process and abolishing invertase by CRISPR/Cas9 genome editing. Frontiers in Bioengineering and Biotechnology. 2020;**8**:1-15. DOI: 10.3389/fbioe.2020.607507

[91] Abdel-Mawgoud AM, Markham KA, Palmer CM, Liu N, Stephanopoulos G, Alper HS. Metabolic engineering in the host Yarrowia lipolytica. Metabolic Engineering. 2018;**50**:192-208. DOI: 10.1016/J.YMBEN.2018.07.016

[92] Madzak C. Yarrowia lipolytica strains and their biotechnological applications: How natural biodiversity and metabolic engineering could contribute to cell factories improvement. Journal of Fungi. 2021;**7**:1-67. DOI: 10.3390/jof7070548

[93] Schwartz C, Curtis N, Löbs AK, Wheeldon I. Multiplexed CRISPR

activation of cryptic sugar metabolism enables Yarrowia Lipolytica growth on cellobiose. Biotechnology Journal. 2018;**13**:1-7. DOI: 10.1002/biot.201700584

[94] Larroude M, Rossignol T, Nicaud JM, Ledesma-Amaro R. Synthetic biology tools for engineering Yarrowia lipolytica. Biotechnology Advances. 2018;**36**:2150-2164. DOI: 10.1016/j.biotechadv.2018.10.004

[95] Wong L, Engel J, Jin E, Holdridge B, Xu P. YaliBricks, a versatile genetic toolkit for streamlined and rapid pathway engineering in Yarrowia lipolytica. Metabolic Engineering Communications. 2017;**5**:68-77. DOI: 10.1016/J.METENO.2017.09.001

[96] Larroude M, Celinska E, Back A, Thomas S, Nicaud JM, Ledesma-Amaro R. A synthetic biology approach to transform Yarrowia lipolytica into a competitive biotechnological producer of β-carotene. Biotechnology and Bioengineering. 2018;**115**:464-472. DOI: 10.1002/bit.26473

[97] Markham KA, Alper HS. Synthetic biology expands the industrial potential of Yarrowia lipolytica. Trends in Biotechnology. 2018;**36**:1085-1095. DOI: 10.1016/j.tibtech.2018.05.004

[98] Madzak C, Tréton B, Blanchin-Roland S. Strong hybrid promoters and integrative expression/secretion vectors for quasi-constitutive expression of heterologous proteins in the yeast Yarrowia lipolytica. Journal of Molecular Microbiology and Biotechnology. 2000;**2**:207-216

[99] Ma J, Gu Y, Marsafari M, Xu P. Synthetic biology, systems biology, and metabolic engineering of Yarrowia lipolytica toward a sustainable biorefinery platform. Journal of Industrial Microbiology & Biotechnology. 2020;**47**:845-862. DOI: 10.1007/s10295-020-02290-8

[100] Larroude M, Park YK, Soudier P, Kubiak M, Nicaud JM, Rossignol T. A modular golden gate toolkit for Yarrowia lipolytica synthetic biology. Microbial Biotechnology. 2019;**12**:1249-1259. DOI: 10.1111/1751-7915.13427

[101] Larroude M, Nicaud JM, Rossignol T. Golden gate multigene assembly method for Yarrowia lipolytica. Methods in Molecular Biology. 2022;**2513**:205-220. DOI: 10.1007/978-1-0716-2399-2_12

[102] Patra P, Das M, Kundu P, Ghosh A. Recent advances in systems and synthetic biology approaches for developing novel cell-factories in non-conventional yeasts. Biotechnology Advances. 2021;**47**:1-29. DOI: 10.1016/j.biotechadv.2021.107695