Biotechnology Advances 35 (2017) 681–710

Contents lists available at ScienceDirect

Biotechnology Advances

journal homepage: www.elsevier.com/locate/biotechadv

Research review paper

Towards systems metabolic engineering in *Pichia pastoris*

Jan-Philipp Schwarzhans[a,b], Tobias Luttermann[a,b], Martina Geier[c], Jörn Kalinowski[b], Karl Friehs[a,*]

[a] Fermentation Engineering, Bielefeld University, Universitätsstr. 25, Bielefeld 33615, Germany
[b] Microbial Genomics and Biotechnology, Center for Biotechnology (CeBiTec), Bielefeld University, Universitätsstr. 27, Bielefeld 33615, Germany
[c] Austrian Center of Industrial Biotechnology (acib), Petersgasse 14, Graz 8010, Austria

ARTICLE INFO

Keywords:
Pichia pastoris
Komagataella phaffii
Non-conventional yeasts
Genetic engineering
Metabolic engineering
Recombinant protein production
Promoters
Systems biology
Physiology

ABSTRACT

The methylotrophic yeast *Pichia pastoris* is firmly established as a host for the production of recombinant proteins, frequently outperforming other heterologous hosts. Already, a sizeable amount of systems biology knowledge has been acquired for this non-conventional yeast. By applying various omics-technologies, productivity features have been thoroughly analyzed and optimized via genetic engineering. However, challenging clonal variability, limited vector repertoire and insufficient genome annotation have hampered further developments. Yet, in the last few years a reinvigorated effort to establish *P. pastoris* as a host for both protein and metabolite production is visible. A variety of compounds from terpenoids to polyketides have been synthesized, often exceeding the productivity of other microbial systems. The clonal variability was systematically investigated and strategies formulated to circumvent untargeted events, thereby streamlining the screening procedure. Promoters with novel regulatory properties were discovered or engineered from existing ones. The genetic tractability was increased via the transfer of popular manipulation and assembly techniques, as well as the creation of new ones. A second generation of sequencing projects culminated in the creation of the second best functionally annotated yeast genome. In combination with landmark physiological insights and increased output of omics-data, a good basis for the creation of refined genome-scale metabolic models was created. The first application of model-based metabolic engineering in *P. pastoris* showcased the potential of this approach. Recent efforts to establish yeast peroxisomes for compartmentalized metabolite synthesis appear to fit ideally with the well-studied high capacity peroxisomal machinery of *P. pastoris*. Here, these recent developments are collected and reviewed with the aim of supporting the establishment of systems metabolic engineering in *P. pastoris*.

1. Introduction

In principle, systems metabolic engineering and systems biotechnology aim to bring the predictability and model based approach of the engineering world to biological systems. For the application of systems metabolic engineering an organism thoroughly studied via omics-technologies, the availability of computational systems biology tools and the capability for targeted genetic engineering, including synthetic biology, is required (Keasling, 2010; Lee et al., 2012; S.Y. Lee et al., 2005; Lee and Kim, 2015). So far, the biotechnological bacterial workhorse *Escherichia coli*, the gram-positive model bacterium *Bacillus subtilis* and the model yeast *Saccharomyces cerevisiae* are the only microorganisms for which systems metabolic engineering is firmly established (Kelwick et al., 2014). Many other organisms show promising features, but not all requirements are fulfilled yet or systems metabolic engineering is currently being established. The very recent progress in *Pichia pastoris* research points towards a growing interest and effort to enable this approach in this yeast. By giving a brief overview of *P.*

Abbreviations: ADH, alcohol dehydrogenase; AOX1/2, alcohol oxidase 1/2; ARS, autonomously replicating sequence; ATF, artificial transcription factor; CDW, cell dry weight; ChiP, chromatin immunoprecipitation; COBRA, constraint-based reconstruction and analysis; CPR, cytochrome P450 reductase; CYP, cytochrome P450 monooxygenase; DAS1/2, dihydroxyacetone synthase isoform 1/2; DDS, dammarenediol-II synthase; DMAPP, dimethylallyl pyrophosphate; DSB, double-strand breaks; GAP, glyceraldehyde-3-phosphate; GCN, gene copy number; GEM, genome-scale metabolic model; GlcNAc, *N*-acetylglucosamine; gRNA, guide RNA; HR, homologous recombination; IPP, isopentenyl pyrophosphate; IR, inverted repeat; MIG1/2, multicopy inhibitor of GAL gene expression 1/2; MIT1, methanol-induced transcription factor 1; MUT, methanol utilization; MVA, mevalonate; MXR1, methanol expression regulator 1; NGS, next generation sequencing; NHEJ, non-homologous end joining; NRG1, negative regulator of glucose-repressed genes 1; OCH1, α-1,6-mannosyltransferase; ORF, open reading frame; PRM1, positive regulator of methanol 1; PTS, peroxisomal targeting sequence; RNA-Seq, RNA-sequencing; ROS, reactive oxygen species; S1,7BP, sedoheptulose-1,7-bisphosphate; UPR, unfolded protein response; VPS, vacuolar protein sorting; WSC, cell Wall integrity and Stress response Component; XYL5P, Xylulose-5-phosphate; ZF, zinc finger

* Corresponding author at: Universitätsstrasse 25, 33615 Bielefeld, Germany.
E-mail address: karl.friehs@uni-bielefeld.de (K. Friehs).

http://dx.doi.org/10.1016/j.biotechadv.2017.07.009
Received 26 May 2017; Received in revised form 20 July 2017; Accepted 24 July 2017
Available online 29 July 2017
0734-9750/

pastoris history, summarizing the newest findings in detail and describing potential future applications, this review intends to aid this line of research.

The methylotrophic, non-conventional budding yeast *P. pastoris* has been established as a wide-spread recombinant protein expression platform in both academia and the industry. According to the web platform www.pichia.com, over 5000 different proteins have been produced in this yeast. The popularity stems from the availability of simple and robust high-cell density cultivation procedures, tightly regulated and extraordinarily strong promoters, good post-translational modification and secretion capabilities, as well as ease of genetic manipulation (Ahmad et al., 2014; Macauley-Patrick et al., 2005). While early ventures focused on technical enzymes (Cereghino and Cregg, 2000), the acquisition of the FDA GRAS (generally regarded as safe) status (Ciofalo et al., 2006) promoted the development of biopharmaceuticals, e.g. the kallikrein inhibitor Kalbitor® or the aglycosylated protease Jetrea® (Corchero et al., 2013; Meehl and Stadheim, 2014). *P. pastoris* has demonstrated its suitability for the expression of targets that proved problematic in other host systems, e.g. membrane bound proteins (Byrne, 2015; Vogl et al., 2014) or glycoproteins (Laukens et al., 2015). Therefore, it has been recommended to consider *P. pastoris* as a standard tool for labs interested in the production of recombinant proteins (Bill, 2014).

Typically, the expression cassette harboring the target gene is integrated into a chromosomal locus via homologous recombination (HR), ensuring a high genetic stability (Cereghino and Cregg, 2000). Alternatively, episomal vectors using the native autonomously replicating sequences (ARS) *PARS1* and *PARS2* are available, but are merely used in a few applications (Cregg et al., 1985; C.C. Lee et al., 2005). While most protein expression studies only require the knock-in of a single target gene, the secretion of fully humanized and terminally sialylated glycoproteins was the largest genetic engineering project in *P. pastoris* to date (Hamilton et al., 2006). This project required deletion of four genes and the integration of 14 foreign genes, including the transfer of the complete human CMP-*N*-actelyneuraminic acid biosynthesis pathway. Other genetic engineering ventures include the introduction of biotin-prototrophy (Gasser et al., 2010), modification of the methanol utilization (MUT) pathway (Krainer et al., 2012) and improvement of protein folding and secretion features (Guerfal et al., 2010). Compared to the wealth of publications focused on protein production and its optimization, relatively few studies concern the biosynthesis of metabolites. The use of non-conventional yeasts in metabolic engineering projects in order to further the yeast platform as a whole was postulated. Only one example from *P. pastoris* was cited (Liu et al., 2013). Nevertheless, chemically different metabolites from riboflavin (Marx et al., 2008) and poly-3-hydroxybutyrate (Poirier et al., 2002) to different carotenoids (Araya-Garay et al., 2012a, 2012b) have been successfully produced in the past.

Although *P. pastoris* shares many properties with the conventional yeast *S. cerevisiae*, it also has its own distinguishing features that provide opportunities or present challenges. Genome integrations are stable but often a high variability of clones from one transformation is encountered, displaying various productivity characteristics or changes in their physiology (Cregg et al., 1985; Schwarzhans et al., 2016a, 2016b). A time-consuming screening process has to be employed in order to find the clone with the optimal features for the desired application (Looser et al., 2015). One of the causes for the clonal variability is the non-homologous end joining (NHEJ) pathway, which mediates integration of foreign DNA at untargeted locations. The clonal variability is named as a key factor holding back the further development of *P. pastoris* as a platform for producing value-added chemicals (Kelwick et al., 2014). On the other hand, *P. pastoris* inherently good protein production features, along with a Crabtree-negative phenotype and a well-developed and well-studied peroxisomal machinery, should make it a desirable host for metabolic engineering projects, surpassing *S. cerevisiae* in certain applications.

The methanol inducible alcohol oxidase 1 (*AOX1*) promoter *pAOX1* and the constitutive glyceraldehyde-3-phosphate dehydrogenase (*GAPDH*) promoter *pGAP* are the most popular choices for facilitating foreign gene expression (Vogl and Glieder, 2013). *pAOX1* offers tight regulation with near-zero transcriptional activity under repressed or derepressed conditions, and exceptionally high activity when induced with methanol (Cereghino and Cregg, 2000). However, the toxicity and flammability of methanol can be of concern as well as the increased oxygen consumption and heat generation of induced high-cell density cultures, necessitating adapted cultivation procedures (Looser et al., 2015). *pGAP* enables constitutive expression at very similar levels to *pAOX1*, reducing process time and handling complexity (Waterham et al., 1997a). However, *pGAP* is not suitable for the expression of host-toxic products, since production and growth phase cannot be decoupled. Variants of both *pAOX1* and *pGAP* with adjusted transcriptional activity have been developed to enable fine-tuned expression experiments (Hartner et al., 2008; Qin et al., 2011). In addition, many new promoters with different regulatory properties and expression strengths have been discovered and applied. They include alternative, methanol inducible promoters of the MUT pathway (Shen et al., 1998; Tschopp et al., 1987; Vogl et al., 2016) as well as other constitutive or repressible promoters (Moreira de Almeida et al., 2005; Stadlmayr et al., 2010).

P. pastoris exhibits a high genetic accessibility. Since its first discovery, scientists have applied various approaches to shape *P. pastoris* towards their needs. While random mutagenesis and subsequent screening procedures were used in the beginning (Liu et al., 1992), selectable markers (Lin Cereghino et al., 2001) and more sophisticated genetic engineering tools like the Cre-Lox recombinase system (Pan et al., 2011) were established over the years. Applying these techniques, genetic and metabolic engineering projects of different scopes have been realized. However, the aforementioned challenges regarding clonal variability and NHEJ off-target integration events complicated genetic engineering projects of higher complexity. Therefore, improvements to the genetic tractability are an ongoing project. In tandem, the repertoire of integrative and episomal vectors requires expansion to facilitate further development of synthetic biology methods in *P. pastoris* (Kelwick et al., 2014).

The commonly used *P. pastoris* strains were genome sequenced between 2009 and 2011 (De Schutter et al., 2009; Küberl et al., 2011; Mattanovich et al., 2009). In the following years, multiple whole genome transcriptomics (Dragosits et al., 2010; Hesketh et al., 2013; Liang et al., 2012), proteomics (Baumann et al., 2010; Dragosits et al., 2009) and metabolomics (Carnicer et al., 2012; Heyland et al., 2011) studies were carried out. In combination with the detailed characterization of the peroxisome (Wriessnegger et al., 2007) and cellular physiology under protein production conditions (Puxbaum et al., 2015), a good basis for *P. pastoris* systems biology was created. Building on these insights, the first generation of genome-scale metabolic models (GEM) was created (Caspeta et al., 2012; Chung et al., 2010; Sohn et al., 2010). However, the incomplete nature of the available genome data and its annotation was cited as a factor holding back systems biology of *P. pastoris* (Dikicioglu et al., 2014).

In this review, we highlight recent advancements to establish *P. pastoris* not only as a recombinant protein production host but as an entire microbial cell factory. Its capability to synthesize various value-added metabolites and to be genetically engineered in a model-based approach have been successfully demonstrated in recent years. The ease of expressing heterologous proteins of bacterial to human origin in *P. pastoris* ought to make it ideal for transferring metabolic pathways from organisms that produce interesting metabolites, but have other disadvantages like difficulties to be cultivated or genetic inaccessibility. In the following sections the newest developments to improve existing advantages of *P. pastoris*, and ways of circumventing drawbacks, are detailed and put into the context of establishing systems metabolic engineering in this non-conventional yeast.

2. Opportunities and challenges of *Pichia pastoris* for metabolic engineering

Although it has been revealed that biotechnologically used *P. pastoris* strains need to be reclassified as *Komagataella phaffii* or *K. pastoris* (Kurtzman, 2009), the original name remains the common designator. From the perspective of metabolic engineering, *P. pastoris* exhibits the same advantages as *S. cerevisiae* when compared to bacterial systems. More complex enzymes can be functionally expressed, viral or bacteriophage infections are not problematic, secretion is more efficient, intra cellular compartmentalization is possible and genome integrations are stable. In addition, *P. pastoris* exhibits certain advantages over *S. cerevisiae*. *P. pastoris* suitability for the expression of various heterologous proteins from bacteria, fungi, algae, plants, animals and humans is well-documented and beyond doubt. Especially recombinant membrane protein expression shows markedly higher success rates in *P. pastoris* than in *S. cerevisiae* (Bill, 2014; Öberg et al., 2011). The hyper-mannosylation is less pronounced, making glycoproteins expressed in *P. pastoris* more likely to be active and less likely to induce hyper-antigenic reactions (Hamilton and Gerngross, 2007). In contrast to *S. cerevisiae*, *P. pastoris* is Crabtree negative, making substrate utilization more efficient under aerobic conditions in shake flask batch cultivations. The peroxisome of *P. pastoris* serves as model for studying varying effects and processes (Waterham et al., 1997b; Wriessnegger et al., 2007). In combination with the recently suggested compartmentalization of metabolic pathways to yeast peroxisomes (DeLoache et al., 2016; Shabbir Hussain et al., 2016), *P. pastoris* distinguishes itself as a very promising candidate for advanced metabolic engineering projects.

Nevertheless, the system also has its drawbacks and challenges. The *S. cerevisiae* research community is one of the largest in the world, enabling considerable scientific output. Due to its smaller size, the *P. pastoris* community exhibits a lower publication rate, although its dedication to better understanding and improving the system by applying the newest technological advances has to be emphasized. Furthermore, the repertoire of genetic tools is not as comprehensively developed for *P. pastoris*. For instance, genome editing and engineering via the CRISPR/Cas9 system was first pioneered in *S. cerevisiae* in 2013 (DiCarlo et al., 2013) and expanded upon in the following years (Generoso et al., 2016; Jakočiūnas et al., 2015; Mans et al., 2015). In the case of *P. pastoris* the method was first established in 2016 (Weninger et al., 2016) and not yet further refined. High cell density cultivation processes are complicated by the high oxygen demand of *P. pastoris* (Looser et al., 2015). But the major challenge holding back more wide-spread application of this yeast is the high clonal variability observed with it (Kelwick et al., 2014).

2.1. Clonal variability

As shown in Fig. 1, clones created during one transformation event can display widely different productivity characteristics or changes in their growth behavior and physiology. This phenomenon has been encountered in early studies with *P. pastoris*, e.g. strains producing the tetanus toxin fragment C exhibited an up to 30 fold difference in product concentration (Clare et al., 1991) and is still prevalent (Cámara et al., 2016; Krainer et al., 2016; Schwarzhans et al., 2016a). However, clonal variability can be beneficial for obtaining clones with high recombinant protein production levels. High producers are usually multicopy strains, an integration event which likely results from in vivo multimerization and subsequent integration of several expression cassettes at once (Aw and Polizzi, 2013; Clare et al., 1991). On the other hand, clonal variability poses a challenge during metabolic and genetic engineering of strains with clearly defined features. Gasser et al. (2010) observed significantly different growth rates of biotin-prototrophy engineered clones from one transformation. It was suspected that variations in the gene copy number (GCN) were responsible. However, in other cases no clear correlation between gene dosage and productivity can be found (Wang et al., 2015). Two strains exhibiting a ca. two-fold difference in product titer did not differ in growth rate or mRNA level of the target gene. In consequence, extensive screening procedures are needed to isolate the transformant with the desired features. Many established screening procedures are available (Looser et al., 2015), with new metabolite-based and automated procedures having been reported lately (Hemmerich et al., 2014; Tredwell et al., 2017). A main culprit for the clonal variability is assumed to be the untargeted integration of expression cassettes at random sites of the genome by the NHEJ pathway (Näätsaari et al., 2012). It mediates the repair of double-strand breaks (DSB) via integration of DNA without requiring homologous sequences, thereby circumventing the targeted HR pathway. This can result in disrupting an untargeted gene and compromising the genetic integrity of the clone. In contrast to *S. cerevisiae*, NHEJ is dominant over HR in *P. pastoris* (Daley et al., 2005). The resulting lower targeting efficiency can be improved, to an extent, by using comparatively long homology sequences of ca. 1000 bp or more (Näätsaari et al., 2012; Nett et al., 2005). It has to be noted, that *S. cerevisiae* has an exceptionally strong HR mechanism. Most other biotechnologically applied microorganisms and cell lines either exhibit a more predominant NHEJ pathway than *P. pastoris* (Guirouilh-Barbat et al., 2004; Meyer et al., 2007; Verbeke et al., 2013) or have such a weak HR pathway that it is essentially not useful for directing foreign gene integration (Belhaj et al., 2015; Rasala and Mayfield, 2015). This issue promoted the recent rise of CRISPR/Cas9 methods which are seen as a promising solution to improve targeting efficiencies, in particular for mammalian and plant cell applications (Belhaj et al., 2015; Komor et al., 2016).

2.2. Ways of controlling clonal variability

If the gene dosage is the key challenge (Fig. 2(A)), the GCN can be determined using proven qPCR or digital droplet PCR methods (Abad et al., 2010a; Cámara et al., 2016). The majority of clones after transformation will contain only a single copy of the gene. Typically, multicopy clones are encountered with a ca. 5–10% frequency and "jackpot" strains (GCN > 10) with about 1% frequency (Aw and Polizzi, 2013; Schwarzhans et al., 2016a). Multiple approaches are possible in order to increase the GCN (reviewed in Aw and Polizzi (2013)). A recent publication presents a post-transformational vector amplification method, aiming to ease the process and make it more cost-efficient by using liquid media procedures (Aw and Polizzi, 2016). Although, fluctuations of productivity between strains can often be traced back to a different gene dosage, there are more possible causes. In addition, a high GCN value can also negatively impact productivity e.g., by triggering the unfolded protein response (UPR) pathway due to elevated stress levels in the cell.

P. pastoris multicopy strains typically contain all copies of the expression cassette adjacent to one another in a tandem array (Clare et al., 1991). Fig. 2 (B) shows how this organization can lead to genetic instability and loss of cassettes due to so called "loop out" events via inter-cassette homologous sequences (Zhu et al., 2009). Using different loci for insertion of multiple copies can mediate the problem but is often not applicable or to work-intense. Instead, optimizing cultivation conditions is advisable in order to find the most stable copy number. Loss of cassettes is often associated with high copy numbers and stress conditions, meaning cells with lower copy numbers have growth benefits. With this in mind, a system has been developed in which only multiple copies of an adenine auxotrophy selection marker with truncated promoter enable survival of the transformant, ensuring constantly high copy numbers (Du et al., 2012). Due to the accumulation of an adenine intermediate in cells with insufficient complementation their colonies appear pink, facilitating direct selection of high copy clones via their white color. Accordingly, the commercially available system has been termed "PichiaPink™".

Due to the location of expression cassettes in a tandem array, their orientation to one another can also have an impact on productivity

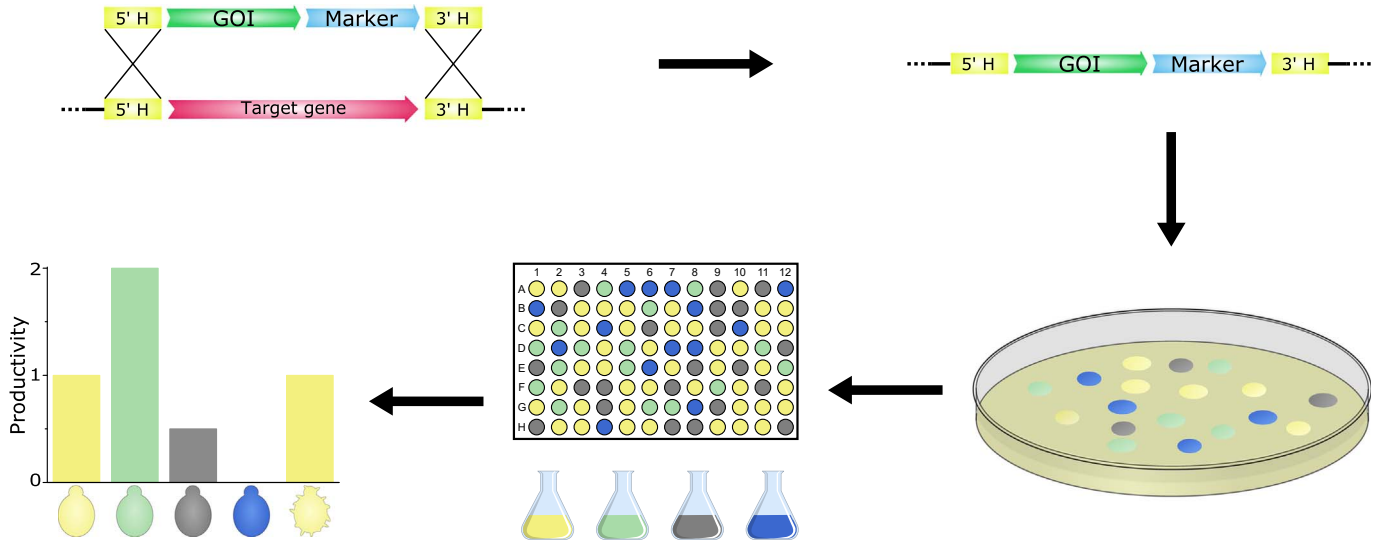

Fig. 1. Common way of integrating the gene of interest in *P. pastoris* and the associated clonal variability. A wide range of productivity as well as physiological changes are encountered in clones from one transformation, making a time consuming screening procedures necessary. 5′ H/3′H = 5′ and 3′ homologous sequence; GOI = Gene of interest.

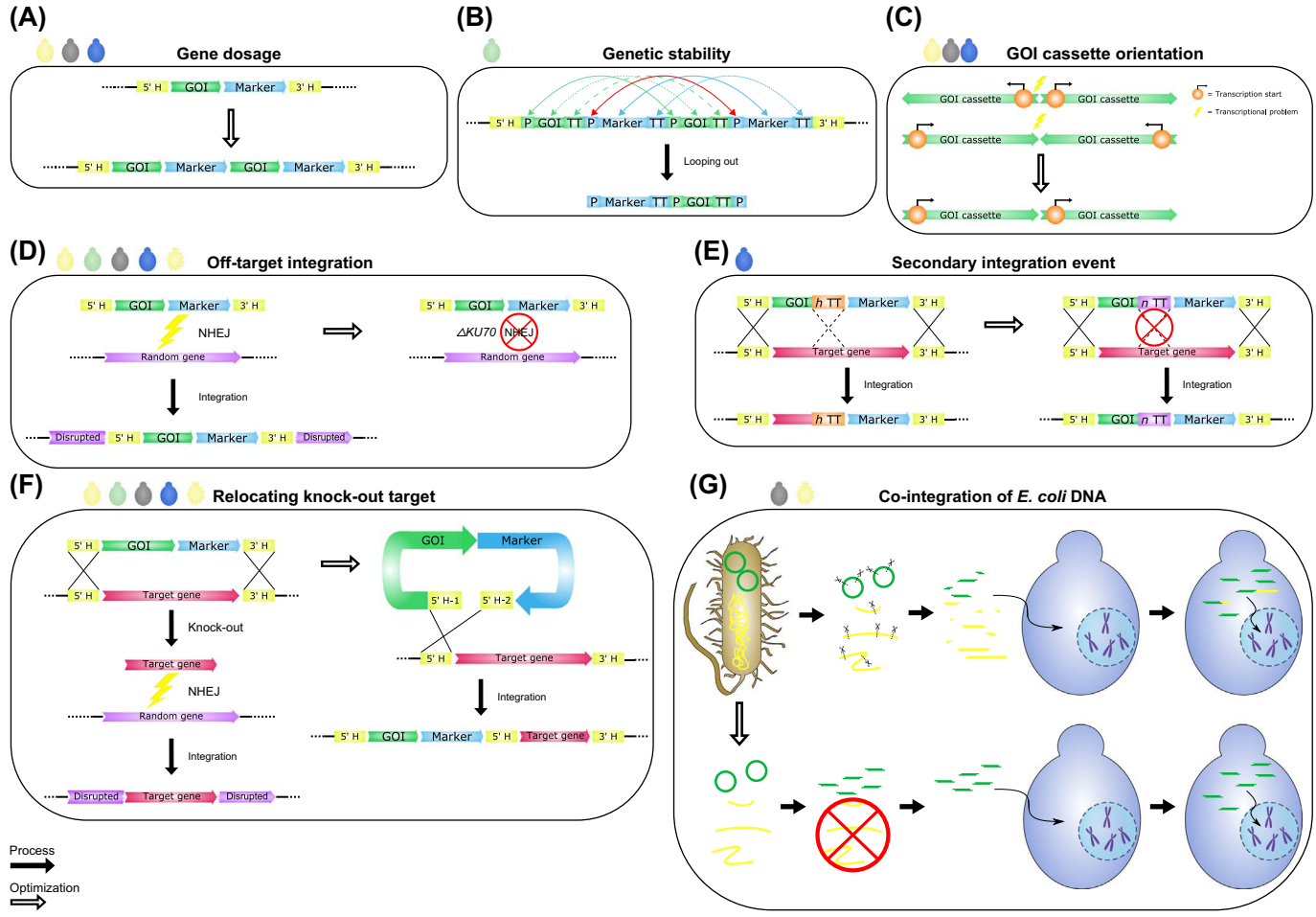

Fig. 2. Possible solutions on the genetic level to solve clonal variability issues. (A) Low gene dosage. Multiple strategies to increase the copy number of the target gene, and thus expression level, are available. (B) Loss of expression cassettes due to homologous sequences between adjacent cassettes. Fermentation procedures have to be optimized in order to minimize loss of cassettes and find a stable copy number. (C) Head-to-head and tail-to-tail cassette organization can be detrimental to productivity due to problems on transcriptional level. It is recommended to target head-to-tail assembly for optimal productivity. (D) Off-target integrations due to NHEJ can lead to unforeseen consequences. Multiple solutions are available, e.g. using a *KU70* deficient strain that is incapable of NHEJ. (E) Integration of only parts of the expression cassette due to internal homologous sequences, e.g. the terminator (*h* TT). This can be solved by replacing them with non-homologous sequence, e.g. a different non-homologous terminator (*n* TT). (F) Knock-out target relocates to different locus via NHEJ and causes same problems as described in (D). By using only an ends-in approach no gene is knocked-out. This approach is not suitable for knock-out studies. Here, screening of transformants must be performed accordingly. (G) DNA from *E. coli* plasmid propagation strain can co-integrate and be actively transcribed in *P. pastoris*. The issue can be circumvented by PCR amplifying the expression cassette.

(Fig. 2(C)). Early it was found, that the head-to-tail organization form is predominant over the head-to-head and tail-to-tail variants (Clare et al., 1991). Recently, a clear correlation between orientation and productivity was discovered via genome sequencing of clones from a library of over 800 strains (Schwarzhans et al., 2016a). While high producer strains contained almost exclusively head-to-tail arrays, mostly head-to-head and tail-to-tail was found in low producing clones. It is suspected, that transcriptional problems cause expression levels markedly lower than the GCN would suggest. In a tail-to-tail organization RNA polymerases from neighboring cassettes run on a converging path, potentially leading to transcriptional arrest due to head-on collision events (Crampton et al., 2006). On the other hand, in the head-to-head organization promoters for the target gene are directly next to one another on adjacent cassettes. This could lead to two RNA polymerases hindering each other from properly starting transcription due to steric interferences. Directed assembly of multiple cassettes into a head-to-tail array prior to integration can circumvent these problems (Vassileva et al., 2001). Interestingly, analyzed strains contained either only head-to-tail or a 50-50 mixture of tail-to-tail and head-to-head arrays. This hints at the existence of two competing integration mechanisms that exclude one another.

As mentioned earlier, NHEJ events can lead to integration of the expression cassette at a random locus on the chromosome (Fig. 2(D)). Depending on the site, genes themselves or 5′ and 3′ regulatory intergenic regions can be disrupted. A preliminary investigation into the distribution of off-target integration events could find no pattern (Näätsaari et al., 2012), but for a conclusive answer a genome wide approach as used for *Kluyveromyces lactis* would be necessary (Kegel et al., 2006). Furthermore, it has been reported for eukaryotic systems that the integration site can influence expression strength of the GOI due to epigenetic factors, chromatin structure and other factors (Day et al., 2000). Using cassettes targeted for different loci, this theory was also investigated in *P. pastoris*. No effect of the integration site on productivity was found (Love et al., 2012; Perez-Pinera et al., 2016). It has to be noted, that loci were chosen deliberately and random NHEJ mediated integration could lead to the disruption of genes of important metabolic pathways that would not be considered in a rationally designed experiment. However, a change of physiological properties due to off-target integration was reported (Schwarzhans et al., 2016b). A gene involved in the oxidative stress response was disrupted, likely causing the change in growth behavior. The strain exhibited a changed colony morphology when grown on agar plates, indicating growth defects. In order to prevent off-target integrations the *P. pastoris KU70* homolog was deleted, thereby eliminating the NHEJ pathway (Näätsaari et al., 2012). Affected cells displayed markedly increased targeting efficiencies and enabled the use of shorter homologous sequences. The generated strain provides a good chassis for further metabolic engineering projects and has been successfully applied to that end (Krainer et al., 2013; Wriessnegger et al., 2014). Nevertheless, the removed NHEJ pathway also lead the cells being more susceptible to DNA damage and ca. 20% lower growth rates, limiting their applicability for industrial purposes. For knock-out studies, a proven alternative presents itself with the split-marker method (Heiss et al., 2013). Here, two fragments of the selection marker are split onto the distal ends of two linear DNA fragments that are transformed into *P. pastoris*. Only simultaneous homologous recombination between both fragments and the target locus leads to integration of the complete selection marker. While transformation frequencies are lower, the approach does not require a specific host strain and is therefore more easily applicable to both academic and industrial research. The method has been applied in various genetic engineering studies (Nocon et al., 2014; Prielhofer et al., 2013; Ruth et al., 2014).

The presence of an additional homologous sequence with the expression cassette can lead to a secondary integration event (Fig. 2 (E)). In the analyzed strains, the event was caused by the *AOX1* terminator (*AOX1* TT) (Schwarzhans et al., 2016a). The chromosomal *AOX1* locus was targeted by two distal homologous sequences (*pAOX1* and 3′ UTR *AOX1* region) on the expression cassette. However, the homologous *AOX1* TT on the expression cassette mediated integration of the selection marker region between *AOX1* TT and 3′ UTR AOX1. The resulting clones pass the selection process but cannot produce any target protein, since they are missing the GOI. They are false-positive. Theoretically, an integration of the GOI region between *pAOX1* and *AOX1* TT without the selection marker is also possible, but clones would not survive the selection process. It was determined that 8% of all > 800 analyzed clones were the result of this secondary integration event (Schwarzhans et al., 2016a). The relatively high frequency of this event means that they pose a burden for the screening process and should be avoided. In order to validate the assumed mechanism, *AOX1* TT was replaced with *CYC1* TT from *S. cerevisiae*. With the non-homologous terminator the secondary integration event was eliminated. However, average productivity was lower, likely because *CYC1* TT was too weak for *pAOX1* (Curran et al., 2013). Many commercial and non-commercial vectors routinely used for *P. pastoris* contain the setup of *pAOX1*, GOI, *AOX1* TT and 3′ UTR *AOX1* (Ahmad et al., 2014; Invitrogen, 2010). Terminators not homologous to the targeted region should replace the *AOX1* TT in such plasmids. To this end, Vogl et al. (2016) assayed various native *P. pastoris* terminators for their effectiveness and highlighted potential replacement candidates. Regulatory elements like *DAS1* TT, displayed effectiveness on the same level as *AOX1* TT and should therefore be suitable for high level production applications.

Another kind of gene disruption was recently discovered. If two homologous sequences are used for targeting the expression cassette (e.g. for knock-out studies), integration results in the excision of the targeted region (Fig. 2(F)). Typically, the excised DNA fragment will be degraded afterwards. In one particular case however, re-integration of the previously deleted *AOX1* locus was observed (Schwarzhans et al., 2016b). Likely mediated by the NHEJ pathway, the knock-out target moved from chromosome 4 to chromosome 2, where it disrupted an untargeted gene during the re-integration event. Only because the disrupted gene belonged to a family of genes involved in signal transduction, membrane trafficking and cytoskeletal organization (Nguyen et al., 2005) the reintegration was discovered. While the relocated knock-out locus was fully functional at its new site, the untargeted gene disruption led to abnormal growth of the affected strain. Therefore, it is hard to gauge how often similar re-integration events occurred that did not result in physiological changes. Again, a genome scale analysis of a large number of clones would be necessary (Kegel et al., 2006). During knock-out studies this re-integration event has to be taken into account, since it cannot fully be avoided. Simply checking strains with PCR assays targeting the knock-out locus might miss its relocation to a new site. In consequence, the effect of gene deletions might be misinterpreted. PCR assays checking for the absence of the knock-out target should complement the established methods. If a knock-out is not essential to the experiment, using only one homologous sequence for integration in an ends-in vector eliminates the possibility of relocating knock-out targets.

Last but not least, the co-integration of *E. coli* DNA from the plasmid propagation strain was detected (Fig. 2(G)) (Schwarzhans et al., 2016b). Common *P. pastoris* techniques use shuttle-vectors for amplifying the plasmid in *E. coli*, extract and digest it, and transform the linear DNA into the yeast cell. During this process contaminating gDNA, and depending on the *E. coli* strain F plasmid DNA, can also be digested and transformed into *P. pastoris*. In vivo ligation capabilities of yeasts in general (Suzuki et al., 1983) and *P. pastoris* in particular (Clare et al., 1991) are well-documented. Based on these abilities, and the integration events discovered during genome sequencing, *P. pastoris* ligated expression cassettes with *E. coli* DNA fragments and integrated the hybrids into its chromosomes. Interestingly, in one strain the integrated *E. coli* DNA was modified after the transformation. Two short palindromic sequences are missing compared to the *E. coli* genome, probably caused by yeast inherent problems with palindromic sequences during

DSB repair (Nag et al., 1989). The integrations were discovered in clones with aberrant colony morphology, which was assumed to be due to the *E. coli* DNA fragments containing genes coding for membrane associated proteins. Using qRT-PCR methods, the presence of mRNA transcripts of the *E. coli* genes could be detected in the *P. pastoris* mutants (Schwarzhans et al., 2016b). However, no proteomic study was carried out, leaving it unclear whether transcripts were indeed translated into proteins. Nevertheless, the co-integration and potential activity of *E. coli* DNA in *P. pastoris* is highly undesirable, especially from an industrial standpoint. The commonly applied gel purification of the expression cassette prior to transformation could mediate some of the observed co-integrations. Yet, integrated *E. coli* DNA fragments ranged in size between 1.5 and 9.3 kb, and could therefore by-pass such a step. Applying PCR for amplification of the expression cassette or the use of rare and blunt-end cutters for digestion should effectively prevent the co-transformation and subsequent co-integration of DNA from the *E. coli* plasmid host.

2.3. Compartmentalization to the peroxisome

Typically, metabolic engineering of microorganisms introduces or modifies metabolic pathways that occur in the cytoplasm. A multitude of native pathways occur simultaneously in the cytoplasm, leading to potential negative crosstalk between targeted and untargeted reactions. This problem severely complicates metabolic engineering projects and can limit achievable product yields. Compartmentalization of engineered pathways to the organelles of eukaryotes presents itself as a good solution to this challenge (Zecchin et al., 2015). Deletion of genes involved in side reaction pathways to reduce crosstalk is a commonly applied tool, but negative impacts on cell growth or the inability to delete essential genes limit this method. In *S. cerevisiae*, localization to the mitochondrion for isobutanol production (Avalos et al., 2013) and vacuolar localization of methyl iodide synthesis have been reported (Bayer et al., 2009). The peroxisome offers distinct advantages over other organelles. Translocation of proteins to the peroxisome can be efficiently realized using peroxisomal targeting sequences (PTS) (Purdue and Lazarow, 2001) and transport of metabolites across the organelle membrane is relatively well understood (Antonenkov and Hiltunen, 2012). Additionally, peroxisome biogenesis can be decoupled from cell growth on certain carbon sources, expanding its use for processes independent of the native cellular machinery (Purdue and Lazarow, 2001). With the β-oxidation of fatty acids, a pathway capable of supplying the building blocks for a variety of value added chemicals is present (Fig. 3) (Poirier et al., 2006).

Recently, three independent studies proposed using the yeast peroxisome for metabolite synthesis (DeLoache et al., 2016; Sheng et al., 2016; Zhou et al., 2016). The product yields for fatty acid derivatives like fatty alcohols and alkenes in *S. cerevisiae* were drastically improved via peroxisomal localization, improving titers more than four-fold (Zhou et al., 2016). A similar approach was used by Sheng et al. (2016) to produce medium chain fatty alcohols, starting from the β-oxidation of fatty acids in the peroxisomes. Both studies clearly show the potential of producing β-oxidation derived metabolites in the yeast peroxisome. DeLoache et al. (2016) work provides a better understanding of peroxisomal membrane transport processes. They improved the common PTS1 tag to lessen the effect of the protein properties on its transport rate. The enhanced PTS1 was used to develop an assay for recombinant protein translocation rates to the peroxisome. Furthermore, the permeability of the peroxisomal membrane was characterized and allowed the construction of a pathway with a permeable substrate but impermeable intermediate. Byproduct formation was significantly reduced and product yields increased.

The peroxisome of *P. pastoris* has been well studied and used as model organelle (Johnson et al., 1999; Liu et al., 1992; Waterham et al., 1997b), making it an ideal target for the proposed compartmentalization of pathways. Specialized isolation protocols have been developed for targeted purification (Wriessnegger et al., 2007). Under inducing conditions, the peroxisome can constitute up to 80% of the total cell volume, indicating a large potential for protein intake (Gleeson and Sudbery, 1988). Peroxisome biogenesis can be induced with methanol or oleate containing medium (Waterham and Cregg, 1997). MUT does not rely on the β-oxidation pathway. Accordingly, down-regulation of β-oxidation pathway genes was observed in cells grown on methanol (Prielhofer et al., 2015; Rußmayer et al., 2015a). On the other hand, glucose limitation led to increased β-oxidation activity. Therefore, fermentation strategies combining glucose growth phase with oleate induced peroxisome biogenesis seem advisable, if high β-oxidation activity is desired. To this end, the characterization of genes involved in oleate-induced peroxisome biogenesis could provide targets for engineering tailored peroxisomes with improved properties for metabolite synthesis (Yan et al., 2008). This approach would also enable a methanol-free cultivation, eliminating negative traits associated with methanol-based processes (Looser et al., 2015). Native (Bhataya et al., 2009; Waterham et al., 1997a) and heterologous (Poirier et al., 2002) PTS have been successfully applied.

Already, the peroxisome has been used for metabolite production in *P. pastoris*. Poirier et al. (2002) demonstrated the peroxisomal synthesis of polyhydroxyalkanoate. Polyhydroxyalkanoate is a 3-hydroxyacyl-CoA polymer, which is an intermediate of the β-oxidation pathway (Fig. 3). Using a whole cell catalyst approach, the enzymatic conversion of cephalosporin C by a D-amino acid oxidase localized at the peroxisome was realized (Abad et al., 2010b). The β-oxidation pathway also feeds acetyl-CoA into the MVA pathway, thereby providing IPP for terpenoid synthesis. This circumstance was exploited by Bhataya et al. (2009) for construction of a lycopene biosynthesis pathway in the peroxisome (Fig. 3). Compared to attempts in which pathway enzymes were localized to the cytosol, four-fold higher lycopene yields were achieved (Araya-Garay et al., 2012b). Along with these proven metabolites, Fig. 3 displays other targets that have been produced in *P. pastoris* and whose biosynthetic pathways could be located to the peroxisome, due to their connection to the β-oxidation of fatty acids. Besides metabolite production, peroxisomal targeting has also been used to express peptides that proved to be toxic for cytosolic expression (Xiao et al., 2016).

By combining recent advances towards establishing yeast peroxisomes as ideal compartments for engineered metabolic pathways and *P. pastoris* naturally well-developed peroxisomal apparatus, metabolic engineering projects in this non-conventional yeast might be considerably boosted. More advanced techniques could enable selective repression of peroxisome biogenesis to reduce undesired metabolic activities in the growth phase. For instance, repression of the peroxisome biogenesis gene *PAS5* via the repressible *pTHI11* promoter should stop peroxisome formation without affecting growth on glucose (Spong and Subramani, 1993).

3. Regulatory elements

To efficiently produce fine chemicals or pharmaceuticals in heterologous expression systems, such as *P. pastoris*, strategies are required to engineer enzyme levels and avoid pathway bottlenecks or resource misdirection (Roquet and Lu, 2014). In eukaryotes, this is often achieved by fine-tuning of transcript levels via employment of different promoters with diverse regulatory and activity profiles. Regarding this, tightly regulated inducible promoters offer certain advantages. They enable decoupling of product formation from cell growth which is a suitable strategy particularly to produce toxic compounds that hamper cell growth. A comprehensive list of established *P. pastoris* promoters and their regulation can be found in the review by Vogl and Glieder (2013). Recently, novel derepressible or inducible promoters were discovered and researchers engineered methanol-free *pAOX1*-based expression systems (Table 1).

Although several promising promoters with interesting features

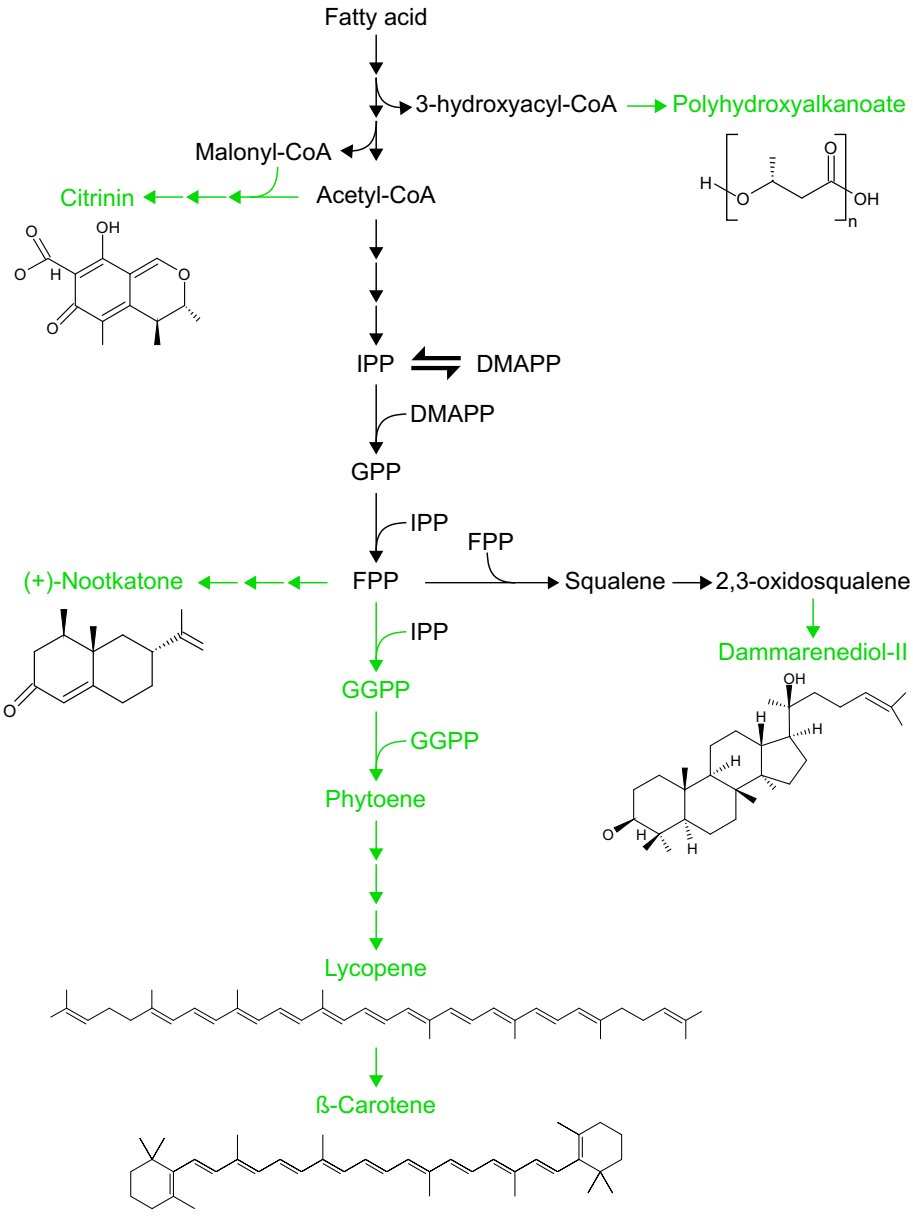

Fig. 3. Examples of heterologous metabolites produced in *Pichia* and their biosynthetic pathway, starting from the β-oxidation of fatty acids. Green arrows/names highlight heterologous pathways and their intermediates or end products. For ease of representation some pathways are shown condensed, indicated by multiple adjacent arrows. DMAPP = dimethylallyl pyrophosphate; FPP = farnesyl pyrophosphate; GPP = geranyl pyrophosphate; IPP = isopentenyl pyrophosphate. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

have been discovered in the recent years, there might be still a variety of novel, differently regulated genes hidden in *P. pastoris*. Typically, DNA microarray analysis has been carried out to identify genes that offer the desired regulatory features (Prielhofer et al., 2013; Vogl et al., 2016). However, since this technique can only give information about relative changes in expression levels, data derived from whole-transcriptome RNA sequencing (RNA-seq) should improve promoter identification. Love et al. (2016) investigated transcriptomic profiles of *P. pastoris* during cultivation on different carbon sources and provide a comprehensive dataset on differently regulated genes for each carbon source. Furthermore, RNA-seq in combination with ribosome profiling offers the possibility to find Kozak sequences that are characterized by high translation efficiencies. Ribosome profiling allows to analyze the translatome by only sequencing the mRNA that has been recovered from translating ribosomes (Ingolia et al., 2009). With this strategy, it is possible to identify genes whose expression is translationally induced under specific conditions, for instance at distinct growth phases (Jeong et al., 2016). Combination of results from both, RNA-seq and ribosome profiling might enable the construction of synthetic promoters in *P. pastoris* that exhibit high transcription as well as translation efficiencies under desired conditions. In a comparable approach, ribosome occupancy was used to fraction mRNA prior to microarray analysis and distinguish between highly- and lowest-translated transcripts (Prielhofer et al., 2015).

3.1. pAOX1 based expression systems

The classic methanol inducible *pAOX1* system remains the most popular one in *P. pastoris*, and it has recently been further developed for methanol-free expression. To this end, it is useful to understand how the MUT pathway and thus aox1 activity is regulated. When *P. pastoris* cells are grown on repressing carbon sources, such as glucose or glycerol, no aox1 activity can be observed (Couderc and Baratti, 1980). This tight regulation is achieved by a variety of activating and repressing transcription factors.

Several trans-acting elements have recently been identified that play a crucial role in the regulatory states of repression, derepression, and induction (Fig. 4(A)). During induction, three transcriptional activators (prm1, mit1, and mxr1) independently bind *pAOX1* at different sites and activate the promotor through a cascade (Wang et al., 2016). Of

Table 1
New promotors and expression systems in *P. pastoris*. The 'modification' column describes genetic changes to the promoter itself or the strain. Approximate levels of protein or gene expression relative to *pAOX1* or *pGAP* are given, where available. Upwards arrows indicate that its corresponding gene is being overexpressed. Abbreviations: mi = methanol-independent; GFP = green fluorescent protein; HRP = horseradish peroxidase; CALB = lipase B from *Candida antarctica*; HAS = human serum albumin; LACB = β-galactosidase; SCP = synthetic core promoter; URS = upstream regulatory sequence.

Name	Regulation	Modification	Expression level	Reference
ADH3	Ethanol induction	Native	Strong (similar to *pAOX1*)	Karaoglan et al., 2016
AOX1	Methanol induction	Native	Strong (naturally ca. 30% of total protein)	Yurimoto et al., (2011)
$AOX1_{mi1}$	Glucose derepression; no Glycerol repression	Δ*NRG1*, Δ*MIG1*, Δ*MIG2*, *mit1*$^{\uparrow}$	77% of *pAOX1* (GFP)	Wang et al. (2017)
$AOX1_{mi2}$	Glucose derepression; no Glycerol repression	Δ*GUT1*, Hp*gcy1*$^{\uparrow}$	25% of *pAOX1* (GFP)	Shen et al. (2016a, 2016b)
$AOX1_{mi3}$	Dihydroxyacetone induction	Δ*DAK1*	50–60% of *pAOX1*	Shen et al. (2016a, 2016b)
$AOX1_{syn}$ (d6)	Glucose/glycerol derepression	Synthetic	400% of *pAOX1* (GFP)	Hartner et al. (2008), Looseret al. (2017)
CAT1	Glucose derepression, methanol/oleate induction	Native	100–180% of p*AOX1* (HRP and CALB)	Vogl et al. (2016)
DAS2	Methanol induction	Native	Strong (similar to *pAOX1*)	Vogl et al. (2016)
FDH1	Methanol induction	Native	Strong (similar to *pAOX1*)	Vogl et al. (2016)
FLD1	Methanol induction	Native	Strong (similar to *pAOX1*)	Shen et al. (1998), Vogl et al. (2016)
G1 (*GTH1*)	Glucose derepression	Native	230% of *pGAP* (HSA)	Prielhofer et al. (2013)
GAP	Constitutive	Native	Strong (similar to *pAOX1*)	Waterham et al. (1997a, 1997b)
LRA3	Rhamnose induction	Native	80% of *pGAP* (LACB)	B. Liu et al. (2016), Y. Liu et al. (2016)
pCore11	Methanol induction	SCP with *AOX1* URS	10% of *pAOX1* (GFP)	Vogl et al. (2014)
pCoreAOX1	Methanol induction	SCP *pAOX1* hybrid variants with *AOX1* URS	35–117% of *pAOX1* (GFP)	Vogl et al. (2014)
Synthetic	Methanol induction/constitutive	De novo SCP variants of *pAOX1*, *pCAT1*, *pDAS1* and *pGAP*	27–122% of *pAOX1* (GFP)	(Portela et al., 2017)
Synthetic	β-Estradiol induction	Synthetic promoter and transcription factor	Stronger than *pAOX1*	Perez-Pinera et al. (2016)
THI11	Thiamine derepression	Native	Strong (similar to *pGAP*)	Stadlmayr et al. (2010), Landes et al. (2016)

these, mxr1 (methanol expression regulator 1) was the first one to be described (Lin-Cereghino et al., 2006). It codes for a protein with a zinc finger binding domain that has high similarity to adr1p from *S. cerevisiae*. When cells are grown on methanol or other gluconeogenic substrates, mxr1 is localized to the nucleus where it binds to sequences upstream of *AOX1*. Cells deficient of mxr1 are unable to utilize methanol, induce transcription of *AOX1* and to form normal-appearing peroxisomes (Lin-Cereghino et al., 2006). In contrast, methanol-induced transcription factor 1 (mit1) does not participate in peroxisome proliferation in response to methanol, but regulates the expression of many genes involved in the MUT pathway (Wang et al., 2016). mit1, together with prm1 (positive regulator of methanol 1), is localized to the nucleus independently of the carbon source. In contrast, mxr1 is only present in the nucleus in the absence of glucose, indicating that mxr1 is involved in derepression of *pAOX1* whereas mit1 and prm1 respond to methanol. The constitutively expressed prm1 binds upstream of *MIT1* and thus induces mit1 expression, if cells are exposed to methanol (Takagi et al., 2012; Wang et al., 2016). All three transcriptional activators are necessary for *pAOX1* activation. However, it remains still unclear how the methanol induction signal is transmitted to prm1 and how it activates prm1 to induce mit1 expression.

In the presence of glucose or glycerol, three trans-acting elements (nrg1, mig1 and mig2) have been recently described to repress *pAOX1* (Wang et al., 2016, 2017). nrg1 (negative regulator of glucose-repressed genes 1) was shown to bind directly to five sites of p*AOX1*. These include two binding sites for the transcriptional activator mxr1. nrg1 deficient strains can express aox1 at low glucose concentrations, indicating that nrg1 competes with mxr1 for *pAOX1* binding sites (Wang et al., 2016). Via live cell imaging of GFP tagged mig1 (multicopy inhibitor of GAL gene expression 1) and mig2, Wang et al. (2017) observed that both repressors are localized to the nucleus when cells were grown on glucose or glycerol. In cells grown on methanol they were predominantly transferred to the cytoplasm. Deletion of these genes led to cells that can express aox1 in the presence of 10 g/L glycerol and the absence of methanol. However, expression of the *AOX1* gene was still repressed if cells were grown on 10 g/L glucose as the sole carbon source. To further elucidate the role of these repressors individually, different deletion strains, including single and double knock-out strains of *MIG1* and *MIG2* were created. A colorimetric aox1 enzyme assay and Western blots revealed that only the double knock-out mutant showed significant aox1 expression, if cells were grown on 10 g/L glycerol. Δ*MIG1* and Δ*MIG2* double knock-out clones showed significantly higher *MIT1* expression levels than the Δ*MIG1* single knock-out strain. Expression levels of prm1 and mxr1 did not change in any of the investigated mutants. This indicates that both mig1 and nrg1 regulate mit1 expression by either binding its promoter or by interacting with the activating transcription factor itself in the presence of glycerol, and that mig2 might increase the repressing effect of mig1 on mit1. Accordingly, overexpression of mit1 from *pGAP* in wild-type cells grown on glycerol leads to minor aox1 activity demonstrating that mig1 and nrg1 rather bind to *pMIT1* than interact with the protein to prevent aox1 activity (Wang et al., 2017). However, no data on mRNA levels of mit1, prm1, and mxr1 if cultivated on 10 g/L glucose was presented. The complete interplay of all known activating and repressing trans-acting elements is not fully understood yet, but the characterization of the aforementioned transcription factors shed light on the regulation of *pAOX1* (Fig. 4(A)).

Based on their characterization results of mig1, mig2 and nrg1, Wang et al. (2017) developed a methanol-independent *pAOX1* expression system via combinatorial engineering of trans-acting elements. Deletion of all three genes resulted in a strain that showed detectable aox1 activity (17% of wild-type cells grown in methanol), if grown on 10 g/L glycerol but was unable to express aox1 in the presence of glucose. Introduction of an additional copy of *MIT1*, constitutively expressed under the control of *pGAP*, let to a further increase of aox1 activity (up to 36% of the wild-type). The corresponding strain will be

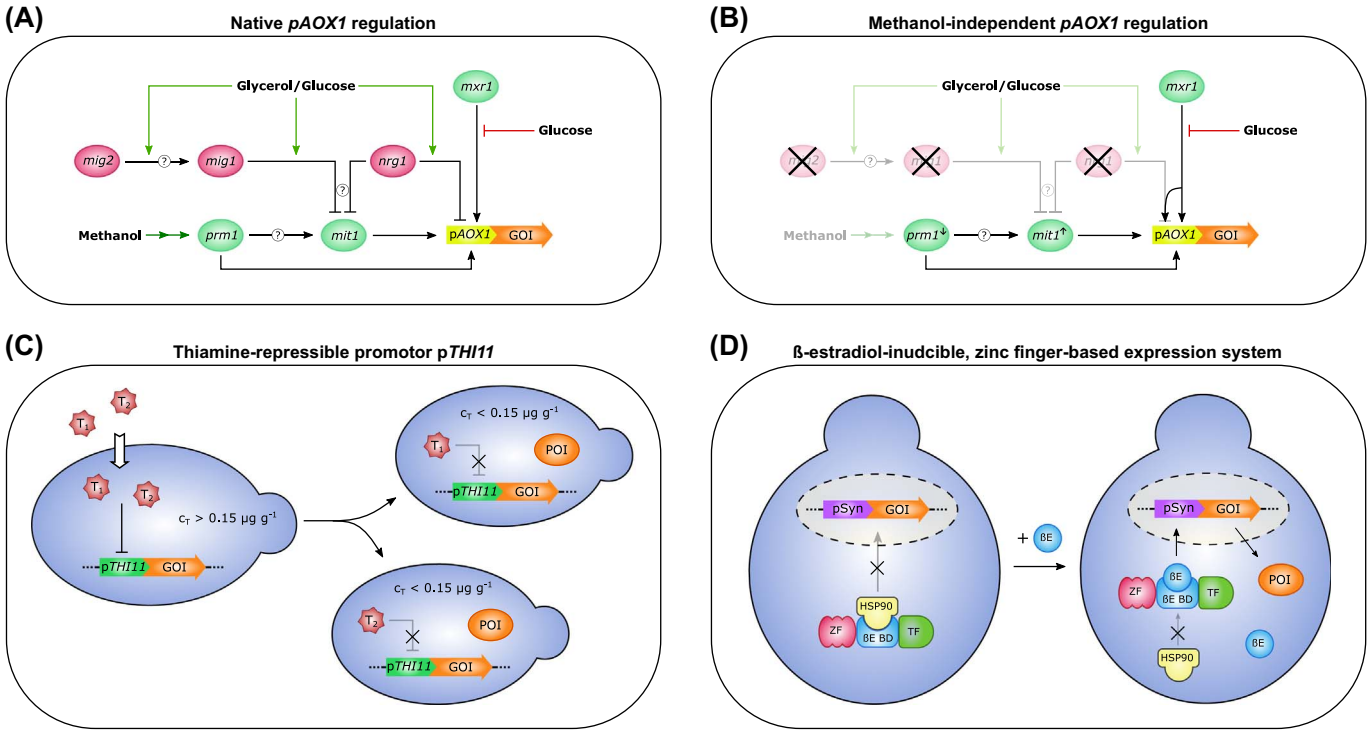

Fig. 4. Regulatory mechanisms of classic and novel *P. pastoris* promoters. (A) The native *pAOX1*. *pAOX1* is induced by methanol and repressed with glycerol or glucose, via a complex and not yet fully understood cascade of transcription factors. (B) Newly engineered methanol-independent *pAOX1*. Via the deletion of three genes (*MIG1*, *MIG2* and *NRG1*) and overexpression of one gene (*MIT1*) the promoter is not activated by methanol anymore. Rather, a glucose repressible expression was achieved. Induction can be realized by shifting the cells from glucose to glycerol medium. The shift enables mxr1 binding to *pAOX1*, with additional sites being available because they are not blocked by nrg1. Due to the absence of methanol, *PRM1* expression is down-regulated. Without the promoting function of prm1, an overexpression of mit1 is necessary to ensure high *pAOX1* activity. (C) Thiamin repressible *pTHI11*. Cells import thiamin (T_1 and T_2) but cannot utilize it. The intracellular content decreases during cultivation due to cell division. Once a threshold of 0.15 µg/g is surpassed, *pTHI11* transcription is initiated. (D) Synthetic expression system. An ATF (consisting of zinc finger (ZF), β-estradiol binding domain (ßE BD) and transcription factor (TF)) is first inactivated via HSP90 and localized to the cytoplasm. Addition of β-estradiol leads to replacement of HSP90 and translocation of the ATF to the nucleus. There, the ZF of the ATF binds to ZF recognition sites and activates the downstream minimal promoter, leading to transcription of the target gene.

referred to as MF1 (Fig. 4(B)). To evaluate the potential of this expression system, GFP was expressed under the control of *pAOX1*. Relative GFP expression in MF1 reached 77 to 130% of the wild-type if grown on 10 g/L glycerol or 0.5% (w/w) methanol, respectively. Differences in aox1 expression and *pAOX1* driven GFP expression in MF1 can be explained by the lower content of peroxisomes in *P. pastoris* cells cultured in glycerol containing media, influencing aox1 content but not recombinant GFP expression. With this system, activation of *pAOX1* transcription can be achieved by derepression from glucose, eliminating the need for methanol induction. Glycerol can be used in the fed-batch phase as the sole carbon source. This allows for high cell density fermentation strategies with auto-induced expression of a recombinant protein via glucose-glycerol shift. However, this engineered expression system also showed detectable GFP expression at low glucose concentrations, indicating that the tight repression of *pAOX1* by glucose was disturbed. This phenomenon might be ascribed to the deletion of *NRG1*, leading to insufficient prevention of *pAOX1* activation by mxr1. Full repression of *pAOX1* in MF1 is only achievable at glucose concentrations of 40 g/L, while low GFP expression could already be observed in the presence of 20 g/L, making this system not suitable for the production of toxic compounds. Nevertheless, this methanol-independent system represents a good alternative to the classic *pAOX1* expression system for regulated gene expression.

A different approach to construct a methanol-free expression system based on *pAOX1* was reported by Shen et al. (2016b). Instead of engineering trans-acting elements to achieve methanol independence of *pAOX1*, they focused on the deletion of kinases to alter or disrupt alternative carbon source metabolism. Their study is built upon the aforementioned kinase single knock-out library (Shen et al., 2016a). The regulation of *pAOX1* depends not only on directly acting, trans-acting elements but also on transporters and sensors that import and recognize carbon sources and thereby initiate complex signaling pathways to either induce or to repress *AOX1* expression. For example, deletion of both the hexose transporter *HXT1* and the hexose sensor *GSS1* led to derepressed expression of *AOX1* in response to glucose depletion (Polupanov et al., 2012; Zhang et al., 2010). However, the exact mechanism of signal transduction from carbon source molecules to the transcription factors is still largely unknown. Shen et al. (2016b) identified the kinases *GUT1* and *DAK*, whose deletion led to methanol-independent *pAOX1* activity. Both kinases are involved in the glycerol metabolism. Their deletion enabled glycerol inducible expression at up to 25% the level of methanol induced *pAOX1*. Despite having a reduced growth rate on glycerol, this system can still present a promising alternative for auto-inducible protein expression when applying a process strategy that comprises a glucose batch phase for biomass generation and a glycerol limited fed batch phase for methanol-free product formation. Induction is facilitated by the glycerol metabolism intermediate dihydroxyacetone (DHA). Supplementation of a Δ*DAK* strain with DHA led to 80–90% of the expression level of the wild-type cells grown on methanol. The approach was validated with different technical enzymes, reaching up to 60% of the native *pAOX1* expression yields.

3.2. Non-pAOX1 based expression systems

Many controllable promoters in *P. pastoris* are derived from genes that are involved in peroxisome biogenesis or the MUT pathway and therefore often are induced by methanol (Vogl and Glieder, 2013). Alternatively regulated promoters are rare. To overcome this limitation, Stadlmayr et al. (2010) identified 24 novel potential regulatory sequences. Of those, the promoter of the thiamine biosynthesis gene

THI11 seemed to be the most promising. The regulation of *pTHI11* is quite interesting since thiamine cannot be utilized by the cells, the biosynthesis of it is very energy-costly, and the import of exogenous thiamine is irreversible (Iwashima et al., 1973; Landes et al., 2016). Landes et al. (2016) were the first to characterize the thiamine-sensitive *pTHI11* and to investigate its suitability for recombinant protein production. In the presence of extracellular thiamine, *P. pastoris* cells rapidly imported the vitamin up to a thiamin level of 1.12 mg/g CDW. During cell growth, the intracellular thiamine level dropped to 0.15 μg/g, indicating that the basal level was reached and maintained by in vivo synthesis. No expression of native and heterologous genes under the control of *pTHI11* was observed, if the intracellular thiamine concentration is above a threshold content of 15 μg/g CDW. After the maximal import of exogenous thiamine, intracellular thiamine levels decrease only via dilution caused through cell division as thiamine cannot be utilized by the cells (Praekelt et al., 1994). Once the thiamine content drops below the basal level, the expression of genes under control of *pTHI11* is enabled (Fig. 4(C)). Under these non-repressing conditions, *pTHI11*-driven protein expression was shown, via chemostat cultivations, to be proportional to cell growth and inversely correlated to intracellular thiamine concentration. Transcript levels of *pTHI11*-controlled genes remained constant at different dilution rates, indicating a constitutive behavior of *pTHI11*. Based on these results, Landes et al. (2016) developed a tailor-made fed batch process strategy to fully exploit the regulatory potential of *pTHI11*. Based on the inability of *P. pastoris* to utilize thiamine and the threshold content needed for repression, they were able to calculate the necessary thiamine concentration in the medium for auto-induction at a desired cell density. After the cells have reached the specific density, a glucose-limited fed batch can be applied to enable a constant growth rate optimal for *pTHI11*-driven expression. The strategy was evaluated against a *pGAP* driven process and similar space time yields were achieved. *pTHI11* regulatory properties represent a promising alternative to *pAOX1*, since it enables programmable, auto-induced protein production. Additionally, *pTHI11* can be used if expression of a recombinant protein needs to be repressed since *P. pastoris* cells take up exogenous thiamine immediately and *pTHI11* responses to intracellular thiamine contents instantaneously (Landes et al., 2016). This strategy was employed by Liu et al. (2015), as discussed in Section 6.1.

To guarantee orthogonality of the promoter, Perez-Pinera et al. (2016) developed a synthetic promoter that is induced by β-estradiol and can exceed *pAOX1* in terms of promoter strength. This promoter consists of the constitutively expressed zinc finger (ZF) DNA binding domain ZF43–8 (Khalil et al., 2012), coupled with the β-estradiol binding domain of the human estrogen receptor. The receptor is expressed in fusion with the VP64 transcriptional activation domain, creating an artificial transcription factor (ATF) (McIsaac et al., 2013). In the absence of β-estradiol, the estrogen receptor interacts with HSP90 which leads to ATF translocation to the cytoplasm (Fliss et al., 2000) and thus prevents the ATF form activating gene expression. If β-estradiol is added to the medium, it is imported by the cells and displaces HSP90 from the estrogen receptor binding pocket. Thereby, the ATF translocates to the nucleus and activates gene expression regulated by a minimal promoter located downstream of multiple ZF binding sites (Fig. 4(D)). To test their ATF, Perez-Pinera et al. (2016) expressed GFP under the control of a minimal *pGAP* and found that expression could be detected with only 0.01 μM β-estradiol. Expression was highest if nine binding sites for ZF43-8 were preceding the promoter. One of the advantages of this system is its highly flexible architecture that enables fine tuning of promoter strength at different levels. These include: promotor driving expression of the ZF, affinity of the ZF DNA binding domain, number of binding sides for the ZF, strength of the transcriptional activation domain and minimal promoter driving expression of the gene of interest. It was found that certain combinations of the ATF regulated, synthetic promoter surpassed p*AOX1* expression levels. Often, these combinations also had greater background expression levels in the absence of an inducer. To characterize their expression system for its suitability to produce multiple proteins induced by different signals, they engineered three different *P. pastoris* strains harboring the ATF to express two different proteins. One target was always expressed under the control of *pAOX1* and the other, as well as the ATF, was regulated by varying promoters. All strains were cultivated in glycerol-containing minimal media for 48 h and could express both proteins selectively at high levels. The induction took place after 48 h of outgrowth to guarantee that *pAOX1* was not repressed by glycerol and could be activated by methanol. Overall, a synthetic β-estradiol-inducible promoter that offers high expression levels and allows for multiplexed expression in *P. pastoris* was developed.

Vogl et al. (2016) carried out a microarray experiment to identify genes involved in the MUT, PPP and ROS (reactive oxygen species) pathways to compare the transcriptional response of *P. pastoris* under glucose-repressed, derepressed, and methanol-induced conditions. They cloned promoters from the aforementioned pathways upstream of GFP to characterize the promoter strength. Many of the identified promoters from MUT genes showed strong methanol-induced activity which was at least half of the *pAOX1* activity. Interestingly, a promoter associated with the MUT pathway displaying high activity (29% of *pGAP*) under derepressed conditions was discovered. The promoter regulates expression of a protein involved in ROS defense and was referred to as *pCAT1*. cat1 likely has functions related to H_2O_2 detoxification arising from ROS stress, which in turn is assumed to activate *pCAT1*. Furthermore, *pCAT1* exhibited tight repression on glucose and could be further induced by methanol, leading to expression levels on par with *pAOX1*. It is also the only promoter of MUT or other related pathways that can be induced by oleate, resulting in comparable expression levels to methanol induction. This feature could proof useful for the aforementioned strategy to compartmentalize metabolic pathways to peroxisomes and use oleate-induced peroxisome biogenesis (Section 2.3). In total, 15 methanol-regulated promoters were identified that offer good options for combinatorial pathway fine-tuning. As a proof of concept, they used a carotenoid biosynthesis pathway for the synthesis of β-carotene, consisting of four enzymes. β-carotene has been produced via the use of a single promoter before (Araya-Garay et al., 2012b). Application of multiple promoters allows for fine-tuning of every single reactions step, reducing the risk of creating bottlenecks in the synthesis of β-carotene. As a result, the β-carotene yield was increased to 5 mg/g CDW, compared to 0.3 mg/g CDW in the original publication. This amount is comparable to optimized *S. cerevisiae* strains (Verwaal et al., 2007). Since different terminator sequences are also beneficial for assembling recombinant pathways, they analyzed a set of 20 endogenous terminators qualified for multiplexed gene expression in *P. pastoris*. All tested terminators showed similar capabilities to *AOX1 TT*, but did not surpass it (Vogl et al., 2016).

Another strategy for coordinated multigene expression involves the usage of bidirectional promoters. Bidirectional promoters facilitate the concurrent transcription of both 5′ and 3′ open reading frames (ORF). They can be found in many eukaryotic organisms, including humans (Trinklein et al., 2004), and are suspected to play an important role in regulatory signal transcription in yeast (Xu et al., 2009). The *GAL1-GAL10* promoter of *S. cerevisiae* is the best characterized native bidirectional promoter in yeast (Johnston and Davis, 1984). It is used in popular vectors as a strong galactose-inducible promoter to drive the transcription of two GOI simultaneously (Partov et al., 2010). Vogl et al. (2015b) published a patent detailing the application of *P. pastoris* bidirectional promoters. The *pHTX1* promoter exhibits a strong constitutive expression profile and consists of a head-to-head fusion of the histone H2A and H2B promoters of *P. pastoris*. Similarly, fusions of *pGAP* and *pTEF* (translation elongation factor 1-α) as well as *pDAS1* and *pDAS2* were used for the creation of constitutive and methanol inducible bidirectional promoters, respectively (Vogl et al., 2015b). These bidirectional promoters were successfully applied for simultaneous transcription of *CAS9* and a gRNA (Weninger et al., 2016), as well as for

expression of a target gene and an UPR pathway transcription factor (Krainer et al., 2016). Their use simplifies vector construction, reduces the risk of loop-out events, and makes the controlled co-expression of targets easier and more predictable.

4. Genetic tractability

P. pastoris exhibits a high genetic accessibility with well-established methods for transformation of various commercial and non-commercial vectors bearing different auxotrophic and antibiotic resistance markers (reviewed in Ahmad et al. (2014). More specialized applications like the aforementioned use of the split marker system (Heiss et al., 2013) and *KU70* deletion strains (Näätsaari et al., 2012), as well as the Cre-Lox recombinase system (Pan et al., 2011) are also available. However, the discussed challenges of clonal variability, NHEJ pathway and lack of episomal vectors complicate genetic engineering projects of higher complexity. To this end, much progress has been made in the last few years to expand the genetic toolbox for *P. pastoris* and streamline the generation of metabolically engineered strains.

4.1. Pathway assembly

In yeast biotechnology genes are typically transcribed in a monocistronic manner, meaning each target gene requires its own promoter and terminator, complicating the construction of larger metabolic pathways. On the other hand, many bacterial operons like the *E. coli* lac operon transcribe a single mRNA for multiple proteins. Such polycistronic operons simplify the necessary vector construction and can enable easier pathway assembly, expression and modification. Using internal ribosome entry sequences, the separate translation of the encoded proteins is realized in bacteria. However, their use is limited by the comparatively high length (~500 bp) and the tendency of downstream genes to be expressed only at low levels (Douin et al., 2004; Nielsen et al., 2009). An emerging alternative are 2A sequences derived from the mRNA translation mechanism of RNA viruses. They are ca. 50 bp short and result in a ribosomal skip during translation, facilitating self-cleavage of the 2A sequence and synthesis of individual proteins from polycistronic mRNA (De Felipe et al., 2006). This co-translational self-processing of multiple proteins under control of one promoter has already been applied in biotechnology (Szymczak et al., 2004). Geier et al. (2015b) transferred and expanded the use of the 2A system in *P. pastoris*. Assaying the suitability of a proven 2A sequence (Shah et al., 2015; Wang et al., 2007) and three variants novel to *P. pastoris*, Geier et al. (2015b) were able to express up to nine genes under control of a single promoter. Furthermore, they successfully tested the possibility of controlling expression strength by placing the genes in different orders, enabling simplified fine-tuning of an entire metabolic pathway. By combining the 2A-mediated polycistronic expression with a bidirectional promoter the simultaneous expression of two different heterologous pathways (β-carotene and violacein synthesis) was realized (Fig. 5 (A)). This approach facilitates the efficient and space-saving construction of pathways in *P. pastoris*. It requires only a single transformation step, contains less internal homologous sequences for looping out events, enables easy tuning of expressional strength and simplifies the regulatory control of expressing multiple pathway genes. Nevertheless, remaining challenges of the 2A system, e.g. incomplete self-processing leading to fusion proteins, have to be addressed in future studies (Geier et al., 2015b).

An essential part to facilitate efficient metabolic engineering of a microorganism, is the possibility to quickly and modularly assemble vectors. This way it is possible to generate large constructs encoding entire pathways by transforming cells with easily synthesizable and transferable small DNA fragments. Assembly can be realized in vitro or in vivo and many techniques like Gibson Assembly, Golden Gate, BioBricks and more have been developed in the last years (reviewed in Casini et al. (2015)). In vivo DNA assembly mediated by complementary overlaps has been widely applied in *S. cerevisiae*, due to the extraordinary HR capability of this yeast (Essani et al., 2015). In contrast, only few reports of its use in *P. pastoris* exist (Yu et al., 2012). Recently, the concept was further developed and characterized. Camattari et al. (2016) demonstrated that overlaps of 20 to 50 bp length were sufficient for efficient assembly of a three-part vector in vivo (Fig. 5 (B)). Molar ratios between the different fragments were optimized, resulting in 90% of transformants containing the correctly assembled vector. The split-marker method (Heiss et al., 2013) was used to increase the frequency of correctly assembled vectors after transformation. Transformation efficiencies were lowered compared to transforming the entire vector at once. Nevertheless, the in vivo assembly enables easy construction of new vector combinations simply by PCR amplification and transformation of the desired components, bypassing the typical *E. coli* cloning step.

On a similar note, an important synthetic biology toolkit from *S. cerevisiae* was transferred to *P. pastoris* very recently. Based on the Golden Gate assembly technique and the MoClo (modular cloning) system of standardized DNA building blocks, a large library of well characterized parts was published for *S. cerevisiae* (Lee et al., 2015). 96 parts are split into categories like promoter, terminator, selection marker and various coding sequences for reporter proteins. All parts were thoroughly characterized for promoter and terminator strength, efficiency of protein degradation tags and targeting efficiency of chromosomal integration, amongst others. To facilitate straight-forward assembly of parts, they contain predefined overhangs compatible with the *BsaI* and *BsmBI* restriction enzymes used for Golden Gate assembly. Using this approach, parts can easily be combined into a single-gene vector, which in-turn can be expanded to a multi-gene vector. The modular nature of the parts, the large number of them and their well characterized features make this tool box ideal for combinatorial metabolic engineering in yeast. Building upon this toolkit, Obst et al. (2017) added *P. pastoris* specific parts (e.g. *pAOX1* and *pGAP*) and characterized them together with selected *S. cerevisiae* parts in *P. pastoris* (Fig. 5 (C)). In total, 26 parts can be combined into 264 different vectors. Addition of the remaining *S. cerevisiae* parts would increase this number to over 4000. While the *P. pastoris* toolkit focuses on (secreted) protein expression, it can easily be extended and applied for metabolite production projects. The publication marks the first *P. pastoris* library of characterized, sequenced and standardized DNA parts for modular vector construction. In order to facilitate efficient distribution to the scientific community, the plasmids bearing the standardized parts are available at the Addgene repository. On a related note, a family of 40 plasmids for use with type IIS restriction enzymes like *BsaI* was reported (Vogl et al., 2015a).

4.2. New integrative and episomal techniques

The CRISPR/Cas9 technology has been discovered and subsequently developed for the application in various organisms over the past years (reviewed in Komor et al. (2016)). Based on a bacterial defense mechanism against foreign DNA invasion, the system was successfully applied to delete or mutate endogenous genes, as well as integrate foreign DNA at a targeted locus. Targeting is facilitated by a ca. 20 bp long gRNA (guide RNA) complementary to the target locus, which directs the cas9 nuclease to the target locus and introduce a strand break (Fig. 5(D)). Breaks can either be repaired by the error-prone NHEJ pathway, resulting in deletions and mutations, or via the HR pathway by providing a homologous DNA repair fragment. Using the HR pathway, foreign DNA can be integrated and due to the DSB break HR efficiency is markedly increased (Storici et al., 2003). In *S. cerevisiae* CRISPR/Cas9 is well-established (DiCarlo et al., 2013) and has been refined for multiplexed-genome editing and transcriptional silencing (Jensen et al., 2017; Mans et al., 2015). Recently, the system was transferred to *P. pastoris* (Weninger et al., 2016). The study is based on the discovery and characterization of multiple nuclear localization

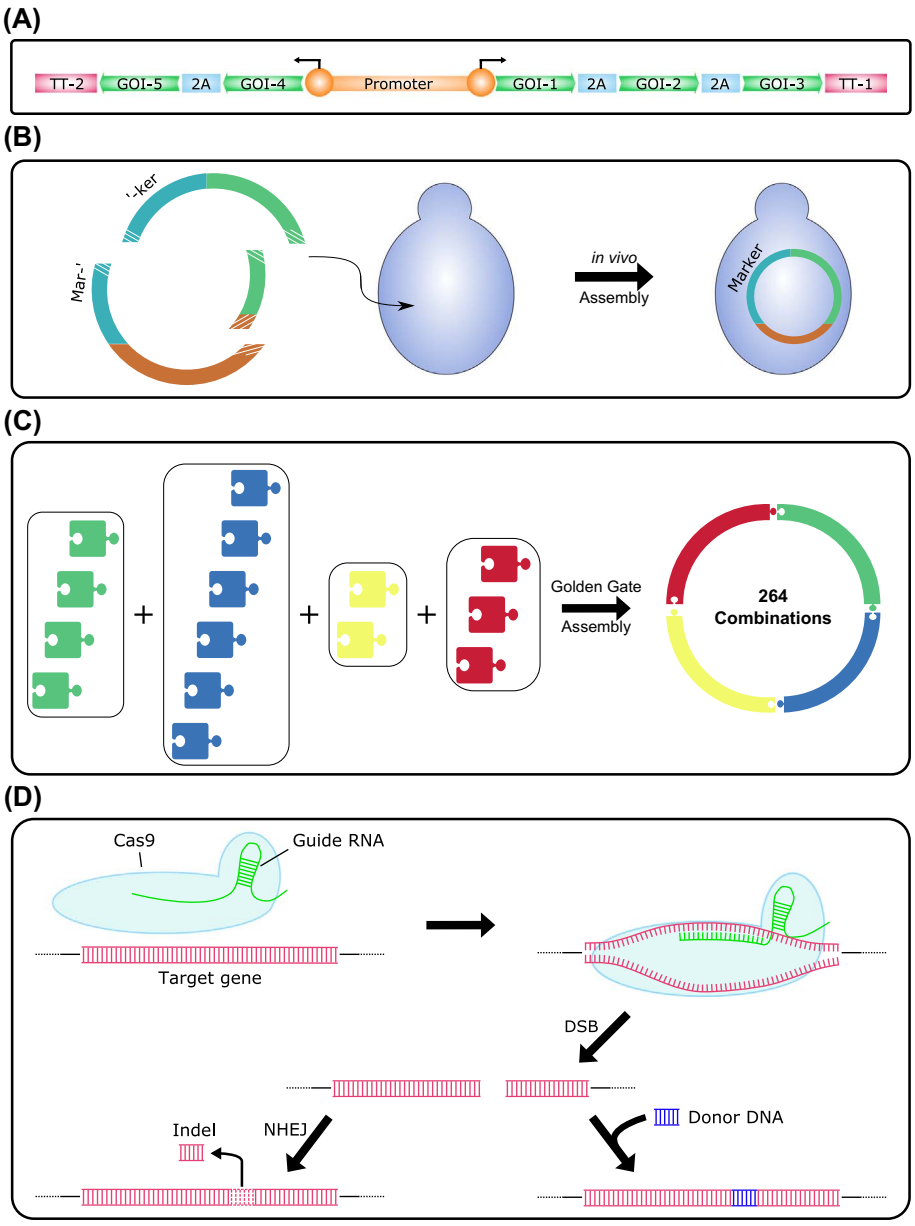

Fig. 5. New tools for genetic manipulation in *P. pastoris*. (A) Combination of bi-directional promoter and 2A-mediated polycistronic expression of a multi-gene pathway. (B) In vivo assembly of an ARS vector via overlapping sequences on the different fragments. Using the split-marker system, the frequency of correctly assembled plasmids is increased. (C) Transfer and adaption of a modular multi-part assembly toolkit to *P. pastoris*. Via Golden Gate assembly a total of 264 plasmid variants can be constructed from standardized and characterized parts. (D) CRISPR/Cas9 cartoon. A DBS is generated and either repaired by NHEJ, leading to indels, or with donor DNA for gene insertion.

sequences necessary for cas9 transfer to the nucleus (Weninger et al., 2015) and bidirectional promoters for co-expression of cas9 and the gRNA (Vogl et al., 2015b). Strikingly, 95 different combinations of cas9 variants (e.g. codon optimized), cas9 and gRNA promoters with and without ribozymes as well as different gRNA sequences needed to be tested to find a functioning setup. Out of these 95 combinations eleven were functional, but only six exhibited targeting efficiencies > 75% with the best one having a 94% efficiency. Cytotoxic effects due to cas9 hyperactivity are suspected to cause the inefficiency of most combinations. Nevertheless, Weninger et al. (2016) demonstrated the highly efficient, targeted and multiplexed genome editing capabilities of their system, combined with an HR integration efficiency on-par with conventional systems. Hence, the currently available CRISPR/Cas9 method should help e.g. in knock-out studies by allowing for time-saving procedures, while further improvements are needed for its application for foreign DNA integration. Potentially, newly developed Cas9 "successors" like the vertebrate based activation-induced cytidine deaminase (AID) tool could help overcome the limitations of the current system. AID was successfully applied in *S. cerevisiae*, where it displayed markedly reduced cytotoxicity (Nishida et al., 2016).

In addition to the established Cre-Lox method (Pan et al., 2011), a new recombinase based system was developed. Three recombinases (BxbI, R4 and Tp901-1) can be used to facilitate the integration of the vector (Perez-Pinera et al., 2016). For targeted integration, a designated "landing pad" of heterologous DNA was previously integrated into the genome of the *P. pastoris* strain. The system offered high targeting efficiencies and via the landing pad an added security measure against off-target activities. To this end it was also applied in the previously mentioned creation of a modular vector toolkit (Obst et al., 2017). It has to be noted, that the use of a landing pad necessitates previous strain engineering work and restricts the potential for future applications. However, an interesting aspect of the presented approach is its capability to be expanded for synthetic gene circuit design. Gene circuits consist of genes flanked by recombinase recognition sites that enable inverting the gene to turn its expression on or off (Roquet et al., 2016). This technology is still in its infancy, but could prove useful for biotechnological applications in the future, as has been demonstrated by creating expression systems that automatically start induction upon reaching a predefined cell-density (Soma and Hanai, 2015).

A different approach for enhanced targeting efficiency was proposed

Table 2
Overview of key genetic engineering projects in *P. pastoris* not targeting the production of metabolites. The number of knocked-out, knocked-in, downregulated and upregulated genes is shown. In case multiple modifications were done separately from one another the number is given as "n × z", with n as the amount of modification targets and z as the number of simultaneous modifications in a strain. Items are ordered chronologically.

Target	Knock-in	Knock-out	Upregulated	Downregulated	Reference
Glycoengineering					
GlcNAcMan5GlcNAc2 N-glycans	3	1	0	0	(Choi et al., 2003)
Man5GlcNAc2 N-glycans	3	1	0	0	(Vervecken et al., 2004)
GlcNAc2Man3GlcNAc2 N-glycans	5	2	0	0	(Bobrowicz et al., 2004)
Fully humanized N-glycans (terminally sialylated)	14	4	0	0	(Hamilton et al., 2006)
Fully humanized N-glycans (nonsialylated)	8	4	0	0	(Li et al., 2006)
Gal2GlcNAc2Man3GlcNAc2 N-glycans	8	1	0	0	(Jacobs et al., 2009)
Strain fitness	0	1	0	0	(Jiang et al., 2015)
N-glycan trimming	3	0	0	0	(Claes et al., 2016)
Recombinant protein production					
Proteolysis	0	2	0	0	(Werten and De Wolf, 2005)
Protein secretion and folding	2	0	0	0	(Gasser et al., 2006)
Protein secretion and folding	0	0	1	0	(Guerfal et al., 2010)
Protein secretion	0	0	8 × 1	0	(Baumann et al., 2011)
Engineered cell cycle phase	0	0	8 × 1	0	(Buchetics et al., 2011)
Methanol utilization	0	0	3	0	(Krainer et al., 2012)
Glycoprotein homogeneity	0	1	0	0	(Krainer et al., 2013)
Central carbon metabolism	0	4 × 1	7 × 1	0	(Nocon et al., 2014)
Heme cofactor availability	0	0	8 × 1	0	(Krainer et al., 2015)
Pentose pathway flux	0	0	4	0	(Nocon et al., 2016)
Protein secretion and proteolysis	0	3	2 × 1	0	(Marsalek et al., 2017)
Other					
Biotin-prototrophy	4	0	0	0	(Gasser et al., 2010)
Peroxisome biogenesis	0	3 × 1	0	0	(Polupanov et al., 2011)
UPR pathway study	1	0	3 × 1	0	(Delic et al., 2012)
Deletion of native secreted protein	0	1	0	0	(Heiss et al., 2013)
NADH regeneration	0	2	0	0	(Geier et al., 2015a)
Fatty acid composition stress response	0	3 × 1	0	0	(Zhang et al., 2015)
Methanol-free AOX1 induction	0	2 × 1	0	0	(Shen et al., 2016b)
Rhamnose utilization pathway	0	1	0	0	(B. Liu et al., 2016)
Kinase involvement in cell growth	0	92 × 1	0	0	(Shen et al., 2016a)
Methanol-free AOX1 induction	0	3	1	0	(Wang et al., 2017)

by Tsakraklides et al. (2015). Cells were reversibly arrested in the S-phase of mitosis, in which DNA is duplicated, via hydroxyurea treatment. It has been shown for *S. cerevisiae* that arrest of cells in this phase led to increased HR activity (Galli and Schiestl, 1996). Similarly, this approach facilitated increased targeting efficiencies in the assayed non-conventional yeasts (Tsakraklides et al., 2015). Especially for *Yarrowia lipolytica* and *Arxula adeninivorans* significant improvements were observed. On the other hand, *P. pastoris* targeting efficiency was only increased from 1.6% to 5.4%. However, very short homologous sequences (50 bp) were used that are not representative of the long sequences of ca. 1000 bp, typically employed in *P. pastoris*. It would be of interest to investigate whether cell cycle synchronization also benefits targeted integration in *P. pastoris* with commonly employed homologous sequences. The system might enable the use of shorter homologs sequences, while enhancing targeting efficiency and not requiring specifically engineered strains.

Besides the clonal variability, the lack of (episomal) plasmids has been cited as a drawback of the *P. pastoris* system (Kelwick et al., 2014). In yeast, ARS are needed for replication of episomal vectors. A genome wide search for ARS consensus sequences resulted in the discovery a large number of potential ARS (Liachko et al., 2014). Some of these were characterized and it was found that, unlike most other yeasts, *P. pastoris* contains mostly GC-rich ARS. In-tandem, an ARS from *K. lactis* (*panARS*) was shown to exhibit a very broad host-range of budding yeasts, including *P. pastoris* (Liachko and Dunham, 2014). These insights were combined to assay the suitability of ARS based vectors for recombinant protein production in *P. pastoris* (Camattari et al., 2016). *panARS* facilitated higher expression levels than the two assayed chromosomal *P. pastoris* ARS and a strain with an integrated expression cassette. Furthermore, the *panARS* vector did not only outperform the productivity of the integrative system, it also provided a much higher clonal homogeneity. Similar observations have been made with an ARS vector system based on a mitochondrial ARS from *P. pastoris* (Schwarzhans et al., 2017). Here, an accidental integration event resulted in the creation of a circular plasmid containing the mitochondrial ARS. Further studies revealed its capability for high level recombinant protein production on-par with *panARS*, lowered clonal variability and suitability for transformation into *S. cerevisiae*. A preliminary scan of the mitochondrial genome revealed the presence of over 500 putative ARS, making them promising candidates for construction of novel *P. pastoris* ARS vectors. Both *panARS* and the mitochondrial ARS significantly expand the episomal plasmid repertoire of *P. pastoris*. Simultaneously, they present an option to tackle clonal variability issue by providing uniform expression levels and causing no genetic perturbation. In consequence, they appear to be best suited for screening experiments, e.g. protein engineering studies. The homogenous productivity would emphasize effects resulting from protein features over those from different expression levels. Both newly reported vectors displayed high stability (close to 100%) under selective conditions, but these values dropped dramatically if the selective pressure was removed (Camattari et al., 2016; Schwarzhans et al., 2017). For auxotrophic markers, a decreasing selective pressure during cultivation, due to release of the relevant metabolite into the medium by prototrophic strains, has to be considered (Pronk, 2002). Suitable selection markers and adapted processes are necessary to ensure stable productivity. Interestingly, for both ARS vectors it was found that productivity was not directly correlated with the GCN (Camattari et al., 2016; Schwarzhans et al., 2017). Strains with markedly different GCN exhibited highly similar protein expression levels. Post-translational and epigenetic factors are suspected to cause this phenomenon (Love et al., 2010). Possibly, *P. pastoris* is capable of higher expression levels with ARS based vectors but an unknown bottleneck is stopping it from reaching the

theoretical potential.

5. Highlights from genetic engineering of platform strains

P. pastoris has been extensively engineered to improve protein production characteristics, study gene or organelle functions and to produce metabolites. To emphasize the increased complexity of these pursuits, the numbers of modified (knocked-in, knocked-out, upregulated and/or downregulated) genes are shown in the following sections. A steady increase in engineering complexity can be noticed, with more studies focusing on simultaneous knock-out and overexpression of genes, instead of constructing separate strains for each modification. First, an overview of key glycoengineering, recombinant protein production and other studies focusing on non-metabolite targets is given. Second, metabolic engineering publications are detailed, showcasing the increased publication frequency in this field since 2014.

5.1. Glycoengineering

The production of glycoproteins received much attention in the first decade of this century (Table 2). *P. pastoris* tends to hyper-mannosylate *N*-glycans with typically around 10–20 mannose residues, displaying a far lower hyper-mannosylation extent than *S. cerevisiae*, in which up to 200 mannose residues can be found (Dean, 1999; Vervecken et al., 2004). This circumstance eased the humanization of the *N*-glycosylation in *P. pastoris*. In 2006, scientists from the company GlycoFi (Lebanon, NH), later acquired by Merck & Co (Boston, MA), succeeded in fully humanizing the N-glycosylation (Hamilton et al., 2006; Li et al., 2006). Additionally, they introduced the full biosynthetic pathway for the human CMP-*N*-actelyneuraminic acid biosynthesis pathway and necessary genes for transfer of the synthesized sialic acid to the mature *N*-glycan. This marked the genetic engineering project of highest complexity in *P. pastoris* to date with 14 knock-ins and 4 knock-outs (Hamilton et al., 2006). A detailed review on the latest developments of glycoengineering of *P. pastoris* can be found in Laukens et al. (2015).

However, after this boom period relatively few publications on the subject were made. One main concern regarding industrial applicability of the engineered strains was the lowered strain fitness. This was in a large part due to the deletion of *OCH1* (α-1,6-mannosyltransferase), resulting in increased flocculation, cell lysis and temperature sensitivity (Davidson et al., 2004; Krainer et al., 2013). Jiang et al. (2015) discovered that deletion of *ATT1*, a homolog of the *S. cerevisiae GAL4* transcriptional activator, improved strain fitness. On the one hand, this result benefits efforts to create glycoengineered *P. pastoris* strains with increased robustness, capable of withstanding the harsh conditions of high-cell density fermentations in large-scale applications. On the other hand, it also furthers our understanding of *P. pastoris* physiological properties. In *S. cerevisiae GAL4* is part of the galactose utilization pathway, but *P. pastoris* is incapable of metabolizing galactose, due to the evolutionary loss of the involved metabolism pathway genes. The conservation of *ATT1* indicates its involvement in the transcriptional regulation of genes unrelated to galactose metabolism. Besides the increased strain fitness, this theory is further corroborated by experiments showing the translocation of att1 to the nucleus and the presence of *ATT1*-associated binding motifs in over 400 promoters of the *P. pastoris* genome (Jiang et al., 2015).

A quite different glycoengineering strategy has recently emerged. Its goal is to completely remove *N*-glycans from proteins that do not require them for their activity and thereby improve product homogeneity. Termed “GlycoDelete”, the system uses a fungal endoglycosidase endoT which removes the entire *N*-glycan (except a single GlcNAc residue) from the glycoprotein when it passes through the golgi (Claes et al., 2016). This approach is especially helpful for the production of membrane proteins, providing a uniform product better suited for pharmaceutical application. Many (viral) membrane proteins can be used as basis for vaccine development, as they often play an important part in receptor-based cell recognition (Fogg et al., 2004). This approach should also have a lower impact on strain fitness than the extensive genetic engineering procedures necessary for humanizing the N-glycosylation. The GlycoDelete system might therefore prove to be of interest for expanding the capabilities of *P. pastoris* for functional expression of glycoproteins whose activity is changed by an incorrect glycosylation, often associated with expression in a heterologous eukaryotic system.

5.2. Recombinant protein production

Besides glycoproteins in particular, the improvement of expression and secretion of heterologous proteins in general has been in the focus of genetic engineering research. Early it was discovered that the co-expression of chaperones and foldases, both heterologous (*S. cerevisiae*) or native, can improve folding and secretion of proteins (Gasser et al., 2006; Guerfal et al., 2010). Endogenous protease genes like *KEX2* and *YPS1* were deleted, thereby reducing proteolysis (Werten and De Wolf, 2005). Often, a specific combination beneficial for one heterologous protein does not show the same effect for another target. This led to the development of several approaches aimed at knocking-in helper proteins and removing proteases to tailor the expression strain to the desired application (reviewed in Puxbaum et al. (2015)). Complex influences of membrane properties (Baumann et al., 2011), the cell cycle (Buchetics et al., 2011) and the MUT pathway (Krainer et al., 2012) on recombinant protein production yields have been assayed in genetic engineering studies by overexpressing selected genes. Typically, the native promoter was replaced with the strong constitutive *pGAP* or the methanol-inducible *pAOX1* promoter.

Historically, genes designated for modification were chosen based on textbook knowledge about metabolic pathways, e.g. genes involved in the ergosterol biosynthesis or methanol metabolism. Nocon et al. (2014) expanded upon this principle by selecting target genes using a model-based approach. Both overexpression and knock-out targets were identified using a GEM by Sohn et al. (2010). Overexpression targets were selected using the Flux Scanning based on Enforced Objective Function (FSEOF) (Choi et al., 2010) and knock-out targets via Minimization Of Metabolic Adjustment (MOMA) (Segrè et al., 2002), respectively. Five out of nine predicted targets resulted in significantly improved protein production. Although many candidate genes were assayed, genetic manipulations were conducted separately with no strain containing more than one modification. It would be of interest to see whether combinations of multiple knock-outs or upregulations result in even better yields and how well these combinatorial effects agree with the model predictions. One of the targets predicted by the models, the pentose phosphate pathway (PPP), was indeed further analyzed and four pathway genes simultaneously overexpressed (Nocon et al., 2016). Some combinations enabled added protein productivity, but others led to unpredicted imbalances in the PPP and lowered product titers.

To increase availability of the heme co-factor for recombinant proteins, the effect of separate overexpression of eight different pathway genes was assayed (Krainer et al., 2015). Interestingly, no increase in heme-requiring protein production was observed. Instead, supplementation of the medium with a heme precursor delivered favourable results. Six more genes involved in the heme biosynthesis pathway were not targeted and no combination of multiple overexpression targets was performed. It is therefore possible, that future endeavors of combinatorial overexpression, potentially based on model predictions, could yield different results.

Besides glycoengineering, a new application for *OCH1* deletions in *P. pastoris* was reported by Weinhandl et al. (2016). Secretion of a recombinant branched chain aminotransferase was improved approximately three-fold by using a *OCH1* deletion strain, compared to the wildtype. Two factors are suspected to facilitate this effect, (i) reduced hyper-mannosylation easing correct exocytosis (ii) increased permeability of the cell wall due to lowered integrity. However, large amounts of native proteins involved in cell wall assembly were secreted

concomitantly, complicating purification of the target protein. It was suspected that inefficient cell wall assembly in growing cells was the root cause. To investigate this thesis, two deletions of genes involved in cell wall assembly were separately introduced in the *OCH1* deficient strain. While secretion rates of the target protein were slightly reduced, the secretion of contaminating proteins was almost completely eliminated in one case. Furthermore, the double knock-out strain displayed reduced cell clumping and a specific growth rate similar to the wild type, in contrast to the ca. 30% decreased growth rate of the *OCH1* deletion strain. These results might also prove valuable for glycoengineering studies, providing more robust and faster growing host strains, similar to the previously mentioned results by Jiang et al. (2015).

Recently, the effect of vacuolar protein sorting (VPS) on protein secretion was investigated by generating VPS mutant strains (Marsalek et al., 2017). The goal was to increase secretion rates of a recombinant protein, tackling a decreased protein yield due to intracellular missorting. While knock-outs of two genes involved in endosomal sorting had a drastic negative effect on secretion rates, a combination with deletion of a cellular protease led to extracellular recombinant protein yields up to 80% higher than in the original strain. Another insight from this study is that *VPS3* is an essential gene in *P. pastoris*, whereas its deletion mutants are viable in *S. cerevisiae* (Peplowska et al., 2007). Together with the diametric opposite effect of deleting VPS genes on protein secretion, a different mechanism of VPS in *P. pastoris* compared to *S. cerevisiae* can be suspected.

5.3. Other genetic engineering studies

Other genetic engineering studies in *P. pastoris* include the construction of a biotin-prototrophic strain (Gasser et al., 2010), studies on peroxisome biogenesis (Polupanov et al., 2011) or the unfolded protein response (UPR) pathway (Delic et al., 2012), as well as removal of the major secreted native protein (Heiss et al., 2013). While some of these studies are closely tied to protein production intents, they also give insights on a more fundamental research level. For example, the biotin-prototrophic strain was presented as being advantageous for protein production processes due to e.g. simpler and cheaper media composition. Nevertheless, the co-expression of four foreign genes in order to construct the biotin synthesis pathway was realized in a more sophisticated way than simply putting all genes under the control of *pGAP*. Integrating results from Stadlmayr et al. (2010) regarding novel native promoters, Gasser et al. (2010) expressed the four heterologous genes under the control of two different promoters. This approach offered the ability to fine-tune expression levels of the pathway genes by using differently strong promoters and also reduced the risk of genetic instabilities due to internal sequence homologies.

In 2015, a chassis strain with markedly improved NADH generation was presented (Geier et al., 2015a). Two genes from the methanol assimilation pathway were deleted, encoding dihydroxyacetone synthase isoform 1 (*DAS1*) and isoform 2 (*DAS2*). Without the gene products of *DAS1* and *DAS2*, the methanol assimilatory pathway is disrupted and methanol cannot be converted into biomass. Instead, methanol is funneled into the dissimilatory pathway, resulting in the generation of two molecules of NADH per molecule of methanol. Geier et al. (2015a) could demonstrate that cells with disrupted methanol assimilation pathway were better suited for a NADH dependent enzymatic whole cell catalysis reaction, if cells were grown in methanol. The constructed strain outperformed the wildtype with an approximately five-fold higher conversion rate. Due to the wide applicability of NADH as co-factor, this approach could benefit many enzymatic conversion projects. The method does seem more suited for whole cell catalysis studies, since the disrupted methanol assimilation pathway results in very low growth rates on methanol. However, by using non-repressing carbon sources in a combined feed the *DAS1 DAS2* double knock-out could potentially also be used in a living cell system (Inan and Meagher, 2001). Thereby, NADH regeneration would be enhanced by the metabolized methanol and e.g. sorbitol used for biomass formation.

In an impressive endeavor, 92 of the 152 annotated protein kinases on the *P. pastoris* genome were separately knocked-out (Shen et al., 2016a). It is suspected, that the remaining 60 kinase genes could not be deleted due to their potential essential functions. The important role kinases play in activating and deactivating proteins via phosphorylation and thereby affecting the cellular machinery on multiple levels is well known. For all knock-out variants the growth behavior on the carbon sources methanol, glucose and glycerol as well as the *pAOX1* activity on the respective media was assayed. Using this approach, 27 strains with impaired growth on specific carbon sources or changes in the regulation of *pAOX1* expression were discovered. This project constitutes the largest single-gene knock-out library in *P. pastoris* to date. Single-gene knock-out libraries, like the Keio collection of *E. coli* with ca. 4000 mutants (Baba et al., 2006), provide an important tool for scientists. In themselves they offer valuable information about the roles of gene products in an organism and serve to predict potential essential genes. Furthermore, they enable researchers to quickly obtain knock-out strains for their experiments and more easily generate multi-gene knock-out strains. Although the library of Shen et al. (2016a) only contains knock-outs for genes of a specific class, it could serve as a starting point for a similar collection of *P. pastoris* single-gene knock-out strains. Deletions were performed in the popular GS115 strain and the basic growth characteristics of all mutants have been characterized and documented, providing a good basis for further expansion.

6. Metabolite production

Over the years a variety of metabolites belonging to different substance classes have been produced in *P. pastoris* (Table 3). However, the metabolite portfolio is by far not as well developed as for recombinant proteins. In the early days of establishing biotechnological applications of *P. pastoris*, metabolite and protein production were simultaneously developed. Researchers from DuPont succeeded in constructing strains producing pyruvic acid (Eisenberg et al., 1997), glyoxylate (Payne et al., 1997), or glyphosate (Gavagan et al., 1997; Payne et al., 1995) and strains capable of stereoselective nitrile hydrolysis (Wu et al., 1999). These achievements are even more impressive due to the lack of knowledge about genomic, transcriptomic and metabolomics properties of *P. pastoris* at that time. Sadly, little progress was made in the following years. Research on establishing *P. pastoris* as protein production host intensified and interest in other projects apparently declined. Nevertheless, in the late 2000s interest seemed to slowly be reignited. Marx et al. (2008) upregulated six genes involved in the biosynthesis of the vitamin riboflavin by replacing their native promoters with *pGAP*. The synthesis of the carotenoids lycopene, β-carotene and astaxanthin was realized (Araya-Garay et al., 2012a, 2012b; Bhataya et al., 2009). Additionally, the hydroxylation of bufuralol to 1′-hydroxybufuralol by a membrane-bound human cytochrome P450 monooxygenase (CYP) in a whole cell catalysis approach demonstrated the possible application of *P. pastoris* as biocatalyst of drug metabolites (Geier et al., 2012). The exceptionally high expression levels of the employed CYP enabled the strain to significantly outperform other assayed microbial hosts during whole cell catalysis, thanks to the good recombinant protein production capabilities of *P. pastoris*. In 2013, the first polyketide of heterologous origin was synthesized in *P. pastoris* (Gao et al., 2013). Requiring the introduction of only two genes, up to 2.2 g/L of the polyketide 6-methylsalicylic acid was produced.

Potentially fueled by the available genome sequences of popular strains and rising amounts of transcriptomics, proteomics and metabolomics data, an increased number of published metabolic engineering studies were visible in recent years (Table 3). In 2014 to 2017, *P. pastoris* was used for the production of compounds belonging to various chemical classes. In the following sections highlights of the recent advancements regarding terpenoids, isoflavonoids, polyketide and other metabolites are discussed.

Table 3
Comprehensive list of genetic engineering projects in *P. pastoris* aimed at producing various metabolites and the number of introduced genetic modifications. Multiple alterations that were done separately from one another are shown as "n × z", with n as the amount of modification targets and z as the number of simultaneous changes in a strain. Entries are ordered chronologically.

Target	Substance class	Knock-in	Knock-out	Upregulated	Downregulated	Reference
Pyruvic acid[1]	Organic acid	1	0	0	0	(Eisenberg et al., 1997)
Glyoxylate	Organic acid	2	0	0	0	(Payne et al., 1997)
Nitrile hydratase activity[1]	Amide	3	0	0	0	(Wu et al., 1999)
Polyhydroxyalkanoate[2]	Polyester	1	0	0	0	(Poirier et al., 2002)
Riboflavin	Vitamin	0	0	6	0	(Marx et al., 2008)
Lycopene[2]	Terpenoid	3	0	0	0	(Bhataya et al., 2009)
Conversion of cephalosporin C[1,2]	Polyketide	1	0	0	0	(Abad et al., 2010b)
Lycopene, β-carotene & astaxanthin	Terpenoid	6	0	0	0	(Araya-Garay et al., 2012a, 2012b)
1′-Hydroxybufuralol[1]	Beta blocker	2	0	0	0	(Geier et al., 2012)
6-Methylsalicylic acid	Polyketide	2	0	0	0	(Gao et al., 2013)
Xylitol	Saccharide	2	0	0	0	(Cheng et al., 2014)
Biodiesel[1]	Fatty acid derivative	1	0	0	0	(Yan et al., 2014)
S-adenosyl-L-methionine	Amino acid	3	0	0	0	(Kant et al., 2014)
Long-chain fatty acids	Fatty acid	2	0	0	0	(Jiang et al., 2014)
Very long-chain fatty acids	Fatty acid	3	0	0	0	(Kim et al., 2014)
Hyaluronic acid	Polysaccharide	2	0	3	0	(Jeong et al., 2014)
Ergosterol	Steroid	1	0	0	0	(Y. Liu et al., 2014)
(+)-nootkatone	Terpenoid	4	0	33 × 1	0	(Wriessnegger et al., 2016, 2014)
Cytochrome P450 reductase stabilization[1]	Terpenoid	2	0	1	0	(Emmerstorfer et al., 2015)
Oosperin	Polyketide	3	0	0	0	(Feng et al., 2015)
Violacein & β-carotene	Polyketide & Terpenoid	9	0	0	0	(Geier et al., 2015b)
Bioethanol	Alcohol	3	0	0	0	(Shin et al., 2015)
Dammarenediol-II	Terpenoid	1	0	1	1	(Liu et al., 2015)
Ricinoleic acid	Fatty acid derivative	2	1	0	0	(Meesapyodsuk et al., 2015)
Fatty acid branched-chain esters	Fatty acid	3	0	0	0	(Tao et al., 2015)
Medium-chain volatile flavour esters[1]	Fatty acid derivative	2	0	0	0	(Zhuang et al., 2015)
Δ9-Tetrahydrocannabinolic acid[1]	Cannabinoid precursor	1	0	0	0	(Zirpel et al., 2015)
6-Hydroxydaidzein & 3′-hydroxygenistein[1]	Isoflavonoid	3	0	0	0	(Chiang et al., 2016; Wang et al., 2015)
Stylopine	Alkaloid	3	0	0	0	(Hori et al., 2016)
Glucaric acid	Organic acid	2	0	0	0	(Y. Liu et al., 2016)
L-Lactic acid	Organic acid	2	0	0	0	(de Lima et al., 2016)
Testosterone[1]	Steroid	2	0	0	0	(Shao et al., 2016)
Dammarenediol-II	Terpenoid	2	0	0	0	(Zhao et al., 2016)
Citrinin	Polyketide	7	0	0	0	(Xue et al., 2017)
Bioethanol[3]	Alcohol	2	0	0	0	(Zhang et al., 2017)

Footnote: 1 = whole cell catalysis; 2 = peroxisomal targeting; 3 = co-cultivation with *S. cerevisiae*.

6.1. Terpenoids

Terpenoids are organic compounds that consist of multiple units of isopentenyl pyrophosphate (IPP) and its isomer dimethylallyl pyrophosphate (DMAPP). Many thousand terpenoid variants are known and a variety of them have proven or potential applications in medicine, cosmetics, food, animal feed, or as platform chemicals (Withers and Keasling, 2007). This led to the exploration of heterologous hosts for producing terpenoids, with one of the all-time success stories being the synthesis of the antimalarial drug precursor artemisinic acid in *S. cerevisiae* (Ro et al., 2006). In *P. pastoris* IPP is generated from Acetyl-CoA via the mevalonate (MVA) pathway (Fig. 3). Unlike bacteria, the MVA pathway is the only IPP synthesis pathway in yeasts and exhibits high efficiencies, resulting in IPP making up 5% of the dry weight under ideal conditions (Lamacka and Sajbidor, 1997). Therefore, they are the ideal host for heterologous production of industrially relevant terpenoids.

The synthesis of (+)-nootkatone was recently developed and improved upon in subsequent publications (Fig. 3). Nootkatone derives its name from its original discovery in the cypress tree *Callitropsis nootkatensis* from the Nootkat region in British Columbia, Canada. It can also be found in citrus fruits like pomelo and grapefruits. A pleasant grapefruit aroma makes nootkatone popular for flavoring, but it also exhibits insect repellent and pharmacological activities such as inhibition of human CYPs. Extraction from fruits is hampered by low yields, whereas chemical synthesis mostly results in the (−)-nootkatone enantiomer with less favourable characteristics (Fraatz et al., 2009). Therefore, biotechnological production of (+)-nootkatone appears to be a promising alternative. Wriessnegger et al. (2014) realized this goal using two different strategies. *P. pastoris* expressing a plant CYP and a cytochrome P450 reductase (CPR) from *Hysocyamus muticus* and *Arabidopsis thaliana*, respectively, was employed for whole-cell catalysis of supplemented (+)-valencene to (+)-nootkatone. A yield of 80 mg/L (+)-nootkatone was achieved. However, problems with the volatility of (+)-valencene complicated this strategy. Therefore, it was decided to enable in situ (+)-valencene synthesis in *P. pastoris* via introduction of a valencene synthase from *C. nootkatensis*. Additionally, in the course of the whole cell catalysis experiments the involvement of a native alcohol dehydrogenase (*ADH*) in (+)-nootkatone synthesis was discovered. Subsequently, the *ADH* was overexpressed. A strain combining heterologous CPR, CYP and valencene synthase expression with overexpression of the endogenous *ADH* was capable of producing ca. 100 mg/L (+)-nootkatone. Lastly, the metabolic flux through the MVA pathway was enhanced by expression of a hydroxy-methylglutaryl-CoA reductase from *S. cerevisiae*. This final step allowed for more than a doubling of the product titer to ca. 208 mg/L. Positive effects of increased flux through the MVA pathway on terpenoid production in *S. cerevisiae* are well documented (Asadollahi et al., 2010). Wriessnegger et al. (2014) demonstrated the applicability of this strategy to *P. pastoris* for the first time. Yields were significantly higher than in *S. cerevisiae*, in which 4 mg/L (+)-nootkatone with added (+)-valencene (Cankar et al., 2014) or 0.14 mg/L with in situ valencene synthesis (Gavira et al., 2013) were achieved. Furthermore, unlike *S. cerevisiae* no cytotoxicity of nootkatone at concentrations > 100 mg/L was observed (Gavira et al., 2013). Building upon these promising results, the *P. pastoris* (+)-nootkatone production system was further optimized in

two subsequent studies. Overexpression of either the endogenous *ice2* or heterologous *S. cerevisae ice2* membrane protein was used to improve CPR stability (Emmerstorfer et al., 2015). While CPR stability was improved, no beneficial effect on (+)-nootkatone production was observed. Based on transcriptomic findings, 33 genes were separately overexpressed to assay whether they benefited (+)-valencene hydroxylation by CYP (Wriessnegger et al., 2016). Interestingly, it was found that overexpression of the gene *RAD52* resulted in two-fold increased productivity. *RAD52* is involved in DNA repair and recombination events in *P. pastoris*, an assumption based on its similarity to the *S. cerevisiae* homolog (Symington, 2002). It is suspected that *RAD52* plays a role in relieving oxidative stress caused by overproduction of terpenoids. Although a direct effect of *RAD52* on CYP and CPR conversion rates could be disproved, no conclusive explanation was found for the increased terpenoid productivity.

The plant terpenoid dammarenediol-II from *Panax ginseng* was produced in a study by Liu et al. (2015). Dammarenediol-II is a cyclized form of 2,3-oxidosqualene, which in turn is formed from the assembly of two squalene molecules (Fig. 3). It can be applied for the synthesis of ginsenosides that are widely used in pharmaceutical applications (Kim et al., 2015). While 2,3-oxidosqualene is a naturally occurring intermediate of the ergosterol pathway in *P. pastoris*, a dammarenediol-II synthase gene (*DDS*) from *P. ginseng* needed to be introduced to enable product formation. This approach resulted in only minute dammarenediol-II yields, compared to high concentrations of ergosterol. It was suspected that diverting the metabolic flux away from ergosterol towards dammarenediol-II should increase productivity. Due to the essential role of ergosterol for cell membrane integrity, a simple knockout of pathway genes was not feasible. Instead, the integration of additional native *ERG1* genes under control of *pAOX1* for increased 2,3-oxidosqualene synthesis was realized. The erg1 catalyzed production of 2,3-oxidosqualene is thought to be the rate-limiting step in ergosterol biosynthesis (Leber et al., 2001). In addition to boosting the precursor pool, flux towards ergosterol was reduced by replacing the native promoter of *ERG7* (catalyzing the first transformation step from 2,3-oxidosqualene towards ergosterol) with the repressible promoter *pTHI11* (Delic et al., 2013; Landes et al., 2016). Applying *pTHI11*, ergosterol synthesis was unaffected during normal growth but could be repressed with the addition of thiamine once cells reached a desired density. To the best of our knowledge, this marks the first published use of downregulated gene expression for metabolic engineering purposes in *P. pastoris*. Compared to the initial attempt with only the heterologous *DDS*, the dammarenediol-II yield was increased from < 0.1 to 0.7 mg/g DCW. Further improvements by upregulating key genes of the MVA pathway are to be expected, as was demonstrated in *S. cerevisiae* (Dai et al., 2013). Dammarenediol-II yields are slightly higher in *P. pastoris* than in *S. cerevisiae* with an unaltered MVA pathway. This suggests that enhancing flux through this pathway in *P. pastoris*, as carried out by Wriessnegger et al. (2014), could result in strains outperforming current *S. cerevisiae* top producers. Another study assayed the impact of co-localizing *DDS* and *ERG1* as fusion proteins or via protein-protein interaction, in order to direct the flux of squalene towards dammarenediol-II (Zhao et al., 2016). Co-localization was realized successfully. However, dammarenediol-II yields of 0.078 mg/g DCW were markedly below the ones achieved by Liu et al. (2015), highlighting the positive impact additional metabolic engineering steps had on productivity.

6.2. Isoflavonoid

Isoflavonoids are a subgroup of flavonoids that can be found in plants and fungi. Over 10,000 different compounds are presumed to exist in nature, fulfilling important roles in plant physiology and antimicrobial activity. Industrial use of flavonoids focuses on dietary supplements, since they have been attributed cognitive preservation effects. Furthermore, their application as part of a personalized medicine (precision medicine) approach to prevent neurodegenerative diseases like Alzheimer's is currently being evaluated (Dixon and Pasinetti, 2010). The soybean isoflavonoids genistein and daidzein have been attributed anti-cancer properties, amongst other potential pharmaceutical features. It is suspected, that the high intake of soy-based food products in Asian countries contributes to the lower occurrence of hormone-related breast cancer, compared to western countries (Barnes, 2010).

Both soybean isoflavonoids have been produced in their hydroxylated form in *P. pastoris*. Regioselective hydroxylation is assumed to enhance bioactivity of the compounds, but these forms cannot be found in nature (Roh et al., 2009). A membrane-bound fusion protein consisting of a *S. cerevisiae* CPR, *Streptomyces avermitilis* CYP and the transmembrane region of an *Aspergillus oryzae* CYP, was used for regioselective conversion of genistein and daidzein to 3′-hydroxygenistein and 6-hydroxydaidzein, respectively (Chiang et al., 2016). This study expands on previous work in which the separate synthesis of both compounds was first described (Chang et al., 2013; Wang et al., 2015). The triple fusion protein facilitated efficient hydroxylation of genistein and daidzein, with a maximal 3′-hydroxygenistein titer of 15.0 mg/L and a 6-hydroxydaidzein titer of 7.5 mg/L. *P. pastoris* clearly outperformed other recombinant host systems, including *E. coli* and *S. cerevisiae*, in both maximal product titer and volumetric productivity. Again, the high aptitude of *P. pastoris* for expression of heterologous membrane bound proteins like CYP and CPR was an important factor towards its application as a host for metabolic engineering. It has to be noted, that production required the supplementation of the corresponding isoflavonoid and a heme precursor, and was realized in the form of whole cell catalysis. CYP utilizes heme as cofactor and therefore, alleviated heme levels benefit productivity. In situ isoflavonoid biosynthesis might enhance yields and could make the process more economical, since addition of external precursors would be avoided. First steps towards reconstructing the necessary biosynthetic pathway in yeast have been made, resulting in *S. cerevisiae* capable of synthesizing various soybean (iso)flavonoids, including genistein, from phenylalanine (Ralston et al., 2005; Trantas et al., 2009). For a fully integrated approach, the overproduction of heme via an augmented metabolic engineering strategy, expanding on the results of Krainer et al. (2015), would be advisable.

6.3. Polyketides

Polyketides are secondary metabolites mostly produced by fungi, plants and bacteria in order to impair growth of competing organisms. Many of them are used as antibiotic agents (e.g. tetracycline and erythromycin) or as anti-cancer drug (e.g. candidaspongiolide) (Donsbach and Rück-Braun, 2005; Staunton and Weissman, 2001). Their biosynthesis starts from malonyl-CoA that can be provided e.g. by the β-oxidation of fatty acids. Naturally polyketide producing microorganisms often exhibit low growth rates, product titers and robustness, complicating process control. Since polyketides can be of very high complexity, multiple enzymes are involved in their synthesis. Therefore, a host proven at expressing various recombinant proteins is advantageous.

Citrinin is a naturally occurring polyketide that can be found in several *Penicillium* and *Aspergillus* species. It acts as antibiotic by inhibiting sterol synthesis but also displays carcinogenic and mutational effects and is classified as mycotoxin (Ostry et al., 2013). Therefore, it was chosen for production in *P. pastoris* solely as well-studied model polyketide compound with a relatively complex structure (Fig. 3). Aided by the discovery and characterization of genes involved in citrinin biosynthesis in a native producer (He and Cox, 2016), Xue et al. (2017) set out to engineer citrinin producing *P. pastoris* cells. Based on their earlier study focused on the synthesis of the polyketide 6-methylsalicylic acid (Gao et al., 2013), a citrinin polyketide synthase and an activating phosphopantetheinyl transferase were introduced into *P.*

pastoris. Fermentation of the recombinant strain led to the formation of a citrinin intermediate but not to the final product. In a stepwise engineering strategy, five heterologous genes from the native producer coding for the (proposed) citrinin biosynthetic pathway were added. In order to minimize the enzyme expression burden on *P. pastoris*, only genes essential to citrinin synthesis were integrated and e.g. associated transporter proteins omitted. This strategy enabled the production of 0.6 mg/L citrinin. Furthermore, a six step synthesis pathway starting from acetyl-CoA and malonyl-CoA condensation was proposed, based on the intermediates discovered during strain engineering steps. All genes were expressed under control of *pAOX1*. Considering the length of the pathway and potential differences in reaction speeds, it is encouraging that citrinin synthesis was successful. Optimizing the expressional strength of selected pathway genes to avoid bottle-necks could improve citrinin yields in future experiments.

The synthesis of the bacterial pigment violacein, a polyketide with potential as both antibiotic and anticancer drug (Choi et al., 2015), was recently reported (Geier et al., 2015b). Five genes from the natural producer *Chromobacterium violaceum* were co-expressed and led to the formation of violacein. Although no product titers were given, the publication underlines the suitability of *P. pastoris* for polyketide synthesis. Starting with the simple polyketide 6-methylsalicylic acid ($C_8H_8O_3$), the complexity was increased to citrinin ($C_{13}H_{14}O_5$) and violacein ($C_{20}H_{13}N_3O_3$). The ease of establishing intricate heterologous pathways, even when expression levels are not optimized, suggests that further development of *P. pastoris* for polyketide synthesis is a promising prospect.

6.4. Other metabolite production studies

Testosterone is a steroid and human sex hormone. It is widely used in medical applications and serves as platform chemical for the synthesis of other pharmacological active steroids (Fernandes et al., 2003). Its biotechnological production has been hampered by side product formation and low conversion rates. Recently, the conversion of a testosterone precursor to testosterone via recombinant *P. pastoris* whole-cell catalysis was reported (Shao et al., 2016). The expression of a single protein from human testis efficiently catalyzed the reaction and no byproducts were detected. However, productivity was further improved by co-expression of the *S. cerevisiae* glucose 6-phosphate dehydrogenase to enhance NADPH regeneration. This step led to a ca. 70% increase in product titer. The beneficial effects of increased NADPH regeneration could be useful for metabolic engineering experiments, akin to the improved NADH regeneration approach discussed earlier (Geier et al., 2015a).

An interesting approach for the production of bio-ethanol with *P. pastoris* was recently realized. Bio-ethanol production from lignocellulosic biomass represents a largely untapped energy source, hindered by the difficulty of fermentation with cellulose, hemicellulose or lignin as carbon source. In order to reduce the need for pretreatment, a co-cultivation system of recombinant *P. pastoris* and *S. cerevisiae* was developed (Zhang et al., 2017). A *P. pastoris* strain expressing two types of recombinant xylanases mediated hemicellulose to xylose conversion. On the other hand, *S. cerevisiae* expressed an *endo-* and exoglucanase as well as a β-glucosidase in order to hydrolyze cellulose to glucose. Using mildly pretreated (0.875% (w/w) H_2O_2, pH 11.5, 35 °C, 1 h) wheat straw as sole carbon source, the ethanol yield was increased to ca. 33 g/L in co-cultivation. In comparison, 25 g/L and 12 g/L ethanol were produced if cultivating *S. cerevisiae* and *P. pastoris* separately, respectively. Complementation of *S. cerevisiae* high ethanol productivity but low xylose utilization with *P. pastoris* high xylose utilization and low ethanol productivity via integrated saccharification and co-fermentation highlighted the possibilities of combining yeasts for metabolite production.

7. Systems biology of *P. pastoris*

The goal of systems biology is to fundamentally understand a biological system in a holistic approach and apply the knowledge to formulate derived mathematical models (Bruggeman and Westerhoff, 2007; Palsson, 2006). By applying these models, desired engineering steps and their outcome can be predicted. Experiments are carried out, the data compared with the predictions and models adjusted accordingly. The systems approach aims to bring the computability and standardized procedures of the mechanical engineering world to biological systems. These do not only benefit basic but also applied research, cutting-down on time necessary for generating strains with desired production characteristics, providing ways to prevent negative engineering effects and enabling deeper insights into hosts for discovery of novel production approaches. An important cornerstone of systems biology is the “omics” technologies, which enable systems-level analysis of the innards of a cell. Genomics, transcriptomics, proteomics and metabolomics are at the forefront when it comes to understanding the composition of an organism, cell or tissue and its reaction to different environmental stimuli. Understanding the physiological properties of the system and connecting them with the molecular level is another essential part of a systems biology approach (Bruggeman and Westerhoff, 2007).

7.1. Omics-based studies

Beginning in 2009, the three commonly applied *P. pastoris* strains GS115 (ATCC 20864) (De Schutter et al., 2009), CBS7435 (ATCC 76273, NRRL Y-11430) (Küberl et al., 2011) and DSMZ 70382 (ATCC 28485, NRRL Y-1603) (Mattanovich et al., 2009) were subject to genome sequencing. These projects primarily relied on 454 next-generation sequencing technology, aided by first generation Illumina methods for the more difficult to resolve sequence parts. Although previous studies applying transcriptomic (Gasser et al., 2007; Graf et al., 2008), proteomic (Dragosits et al., 2009) and metabolomic (Solà et al., 2007) assays had been performed, the availability of genome data allowed for a much deeper understanding of cellular processes. Pre-genomic transcriptomic projects often relied either on heterologous microarrays with *S. cerevisiae* probes or on early designs for *P. pastoris* based on limited insights into the genome. Thanks to the genome sequence data, subsequent endeavors could more easily connect observations between different omics-levels. In the following years, studies involving omics-technologies were published at an increasing rate (Fig. 6). Milestones include the first application of genome-scale RNA-Seq, focused on the transcriptional response to growth on different carbon sources and improved annotation (Liang et al., 2012). Methods for metabolic sampling were developed (Tredwell et al., 2011) and metabolic flux analysis revealed the interplay of the central carbon metabolism (Jordà et al., 2013; Srivastava et al., 2012). Especially recombinant protein production associated characteristics, like the UPR pathway (Dragosits et al., 2010; Hesketh et al., 2013) along with production and secretion capabilities (Baumann et al., 2010, 2011, Pfeffer et al., 2011, 2012) were in the focus of systems level experiments. However, recent advancements might enable not only a deeper understanding of *P. pastoris*, but also provide a fundament upon which to build a truly holistic model of *P. pastoris* cellular processes.

Owing to the limits of the technology available at the time, the first *P. pastoris* genome sequencing projects were not able to resolve the entire genome. Multiple gaps (> 10) of up to 6 kb were left open. The last years have seen the rise of “next generation sequencing” (NGS) technology, facilitating higher coverage, sequencing depths and resolution of previously difficult to sequence regions like homopolymers, while costs dropped dramatically (Van Dijk et al., 2014). Recently, multiple studies focused on improving the available genome data by applying the latest NGS technology (Table 4). Love et al. (2016) simultaneously assayed GS115, CBS7435 and DMSZ 70382 via PacBio

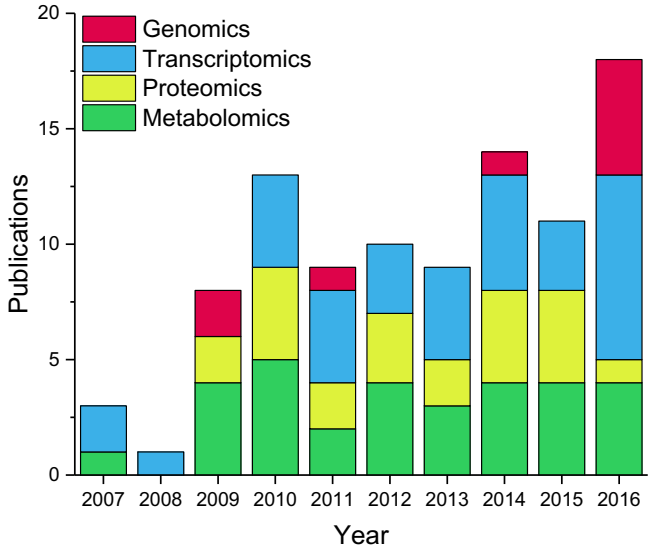

Fig. 6. Omic studies in *P. pastoris* between 2007 and 2016. A first spike at 2009–2010 is visible, when the first genome sequences were published. The second, larger, spike in 2016 is likely due to a combination of multiple studies focused on improving genome data and utilizing transcriptomic techniques.

SMRT and Illumina HiSeq2000 systems. In this comparative approach, they were able to close all remaining gaps. The experiment was complemented by RNA-Seq analysis of cells grown in glucose, glycerol and methanol medium with four time points during a 48 h cultivation. Differences in transcriptional regulation between strains and sampling points were elucidated, leading to the discovery of multiple genes with interesting transcription patterns and strengths. They were proposed as candidates from which novel *P. pastoris* promoters can be derived. The accumulated data was also used to update the nomenclature and determine > 150 new alternative splicing variants. However, the work was focused on the comparative aspect of the three analyzed strains and the implications of the discovered differences on both genome and transcriptome level. Two additional studies published that year made it their prime objective to improve the *P. pastoris* CBS7435 genome data and provide a highly resolved, characterized and annotated basis for other scientists to use as basis for their systems level experiments. While Sturmberger et al. (2016) applied PacBio SMRT and Illumina HiSeq2500 NGS technology, Valli et al. (2016) analyzed RNA-Seq results of 37 different conditions. Both studies were complemented with available proteomic data (Renuse et al., 2014; Rußmayer et al., 2015a). Amongst their key genomic results are the closing of all remaining gaps, identification of two linear plasmids as well as resolving the telomeric and ribosomal repeats. The transcriptomic data was used to manually curate nearly all ORF (5111 of 5256), detect almost 500 new ORF (ca. 150 confirmed via proteomics data), identify ca. 5000 putative UTR and detect a large number of alternative translation start sites (ca. 700) and upstream ORF (ca. 900). Important elements involved in post-transcriptional (UTR, upstream ORF) and translational (alternative translation start sites) regulatory elements were characterized. Furthermore, the *P. pastoris* gene nomenclature was updated genome-wide using data for *S. cerevisiae* homologs and available *P. pastoris* experimental data. Genes without homologs or with uncharacterized homologs were annotated with their hypothetical function based on e.g. BLASTp results. In addition to the updated annotation, gene names were standardized according to the established three letter plus number code system. Via this systematic annotation and renaming process the second best functionally annotated yeast genome was generated.

In order to provide an easy access to the second generation *P. pastoris* genome data, the website www.pichiagenome.org was updated with the new genome, annotation and nomenclature data (Sturmberger et al., 2016; Valli et al., 2016). By modeling the nomenclature and annotation data on *S. cerevisiae*, the transferability of methods and results between these two yeasts will be significantly improved. Researchers can rely on a central database with comprehensive as well as exact information on genome sequence, annotation of ORFs, predicted protein domains and localizations as well as available *S. cerevisiae* homolog data. In comparison to its "bigger brother", the *S. cerevisiae* database at www.yeastgenome.org, additional features like metabolic pathway information are currently missing. But the current development of www.pichiagenome.org signals an increasing interest of its community to create a similar service.

Many studies in the last years applied different omics-technologies to interrogate *P. pastoris*. Transcriptomic analysis has been employed to elucidate the effect of growth rate on protein production and secretion, stress response, mating and maintenance energy demands (Rebnegger et al., 2014, 2016). Furthermore, a transcription factor with potential benefits for recombinant protein production was characterized (Ruth et al., 2014) and the tendency of *P. pastoris* to regulate its expression on the transcriptional rather than the translational level revealed (Prielhofer et al., 2015). Metabolomic studies focused on the ^{13}C method, in order to analyze fluxes through the central carbon metabolism under different conditions. To this end, Jordà et al. (2013, 2014) applied instationary ^{13}C-metabolic flux to improve upon previous NMR

Table 4
Recent studies concerning the advancement of *P. pastoris* genome data.

Project	Strains	Genomics	Transcriptomics	Proteomics	Key findings
(Love et al., 2016)	GS115, CBS7435, DSM70382	PacBio SMRT & Illumina HiSeq2000	10 conditions	–	– All gaps closed – Linear plasmid – Nomenclature updated – 196 (putative) alternative splicing variants – Comparative transcriptomics
(Sturmberger et al., 2016)	CBS7435	PacBio SMRT & Illumina HiSeq2500	2 conditions	(Renuse et al., 2014)	– All gaps closed – Two linear plasmids – Telomeric region and ribosomal repeats resolved – 5111 ORF manually curated – Putative chromosomal centromere
(Valli et al., 2016)	CBS7435	–	37 conditions	(Rußmayer et al., 2015a)	– Nomenclature updated – 492 new ORF; 152 confirmed via proteome – 4916 putative UTR – 341 ORF predictions corrected – 659/885 genes with alternative start site/uORF

data. Valuable insights into how *P. pastoris* adjusts important metabolic pathways (including glycolysis, PPP, TCA and methanol oxidation) under different conditions like recombinant protein production or changing carbon sources were obtained. A combination of metabolomic and transcriptomic methods led to the discovery of stress induced arrest of growth and recombinant protein production following methanol induction (Edwards-Jones et al., 2015). The observed effect was reversible, leading to a delayed production of the target protein and temporary growth reduction. Interestingly, the results indicated no clear correlation of the phenomena to the UPR pathway, suggesting involvement of other factors like the redox state of the cell. An optimized metabolome sampling technique was developed, reducing the loss of low molecular weight compounds during cell separation (Rußmayer et al., 2015b). The spectrum of measurable metabolites in ^{13}C-flux analysis was considerably expanded by the establishment of GC-Quadrupole Time-of-Flight Mass Spectrometry (GC-QTOFMS) procedures for *P. pastoris* experiments (Mairinger et al., 2015).

The multi-omics study of the greatest scope so far for *P. pastoris* analyzed the methanol metabolism on transcriptomic, proteomic, and on metabolomic level (Rußmayer et al., 2015a). Genome-wide transcription analysis was combined with the measurement of 575 proteins, 141 metabolite pools and 39 fluxes involved in MUT under different carbon source regiments. Supporting previous findings (Prielhofer et al., 2015), a strong positive correlation between transcript abundance and protein concentration was found. A key finding of the study was the presence of all enzymes of the MUT pathway in the peroxisome. Previously, it was assumed that parts of the pathway take advantage of the cytosolic PPP for recycling of xylulose-5-phosphate (XYL5P). However, Rußmayer et al. (2015a) discovered isoforms of PPP enzymes containing PTS1 signal sequences that are involved in XYL5P recycling. The prevalence of peroxisomal localization of methanol induced XYL5P recycling was further corroborated by increased transcription for PTS1 containing isoforms in the presence of methanol, while those without were downregulated. Isolation of peroxisome organelles and subsequent proteome analysis validated that PTS1 isoforms were present in methanol induced cells and absent in non-induced cells. In addition to its peroxisomal localization, a novel rearrangement in the XYL5P recycling pathway was discovered. The XYL5P intermediate sedoheptulose-1,7-bisphosphate (S1,7BP), along with the associated transcript and enzyme, were found in methanol grown cells. S1,7BP is an important part of the Calvin cycle in chloroplasts (Raines et al., 1999), but its presence in *S. cerevisiae* has also been shown (Clasquin et al., 2011). Via S1,7BP and the aforementioned PPP enzyme isoforms, the in-organelle generation and regeneration of XYL5P from methanol is enabled. This eliminates the need for XYL5P import from the cytosol, as proposed by previous models of MUT in methylotrophic yeasts (van der Klei et al., 2006). A model was postulated, equating peroxisomal compartmentalization of MUT in methylotrophic yeast to the carbon assimilation pathways of phototrophic organisms to their chloroplasts. The evidence presented supports this thesis and additional experiments with strains deficient in key peroxisomal isoforms could further cement its validity. Due to the high expression levels of proteins for MUT, the synthesis of related cofactors was also increased. Particularly the observed regulatory effects on genes involved in heme biosynthesis could serve as basis for optimizing previous attempts at increasing heme production in *P. pastoris* via metabolic engineering (Krainer et al., 2015).

7.2. Genome-scale modeling

Systems biology relies upon the use of data gained from high throughput omics-technologies for the creation of genome-scale metabolic models (GEM) (Bruggeman and Westerhoff, 2007; Price et al., 2004). In particular for biotechnological applications, the predictability that comes with having modeled central cellular processes is beneficial (Otero and Nielsen, 2010). Model-based approaches have been successfully used in several industrial microorganisms to purposefully predict genetic engineering targets for enhanced product yields and improved cultivation characteristics. Constraint-based reconstruction and analysis (COBRA) methods have proven their suitability for such applications (King et al., 2015; Schellenberger et al., 2011). Shortly after the first genome projects were published, GEM for *P. pastoris* followed. For *P. pastoris* DSMZ 70382 the model *PpaMBEL1254* (Sohn et al., 2010) and for strain GS115 the model *iPP668* (Chung et al., 2010) were published. They are of similar size with ca. 1300 metabolic reactions involving ca. 1200 metabolites, segregated into 8 compartments. Two years later, a refined GS115 model (*iLC915*) introduced simultaneous computation of recombinant protein production and MUT, as well as optimized predictability of growth behavior on methanol based media (Caspeta et al., 2012). However, *iLC915* omitted extracellular and nuclear reactions of *iPP668*. Notably, the gene coverage was increased from 10% (*PpaMBEL1254*) to 17% (*iLC915*). Compared to current *S. cerevisiae* GEM, less reactions but a higher percentage of genes are covered (Österlund et al., 2013; Sánchez and Nielsen, 2015). Recently, the established *P. pastoris* GEM were further developed and also applied in metabolic engineering projects.

iLC915 was refined to include both native *P. pastoris* and humanized N-glycosylation of glycoproteins (Irani et al., 2016). Integration of *N*-glycosylation required reconstruction of the dolichol pathway for native glycosylation. Hyper-mannosylation was set to 9 mannose residues at the lower end of the observed range (Vervecken et al., 2004). The necessary pathways for synthesis of fully humanized and terminally sialylated N-glycosylation were taken from literature (Hamilton et al., 2006). Subsequently, the GEM was named *ihGlycopastoris*. For validation, the GEM was evaluated against published *P. pastoris* data. Both growth rate and recombinant protein productivity were assayed under various cultivation conditions and with multiple target proteins. *ihGlycopastoris* displayed excellent consensus ($R^2 = 0.92$) for the growth rate, but failed at predicting recombinant protein production rates. Protein production rates were often considerably overestimated by the model. It is suspected that stress factors like UPR, secretion capacity and others (Puxbaum et al., 2015) that are not considered by *ihGlycopastoris* lead to the overestimation of production rates. Despite the aforementioned challenges, the model is to be regarded as the first step towards a fully realized N-glycosylation GEM. Experimental data on proteome-wide native glycosylation in *P. pastoris*, as was done for *S. cerevisiae* (Chen et al., 2014), and implementation of the effects of various stress factors on recombinant protein production rates are necessary to further develop *ihGlycopastoris*. Nevertheless, the results are a clear indication of the burden on *P. pastoris* metabolism due to glycoprotein expression. The model highlights the need for expanding current GEM to increase their applicability and make them suitable for a wider range of tasks.

In a different approach, previously published GEM were combined and expanded upon into a new “meta”-GEM named *iMT1026* (Tomàs-Gamisans et al., 2016). *iPP668*, *PpaMBEL1254* and *iLC915* were used for assembly of *iMT1026*. *ihGlycopastoris* was published later and could therefore not be considered. In addition to combining the three previous GEM, existing pathways in the models were reevaluated based on published data. For example, carbon source utilization pathways that were originally transferred from *S. cerevisiae* models, but for which the non-utilization in *P. pastoris* was proven, were removed. On the other hand, the fatty acid metabolism and oxidative phosphorylation required the addition of new reactions or condensation of existing ones. In the end, 185 new reactions were added, based on literature data. The resulting *iMT1026* exhibits the highest gene coverage (ca. 19%) as well as number of metabolites and reactions (1689 and 2035, respectively) of any *P. pastoris* GEM published so far. The validity of its prediction capability was confirmed via comparison to published data. *iMT1026* growth rate, CO_2 production and O_2 consumption simulations were in very good agreement with experimental data (< 10% deviation), outperforming *iPP668*, *PpaMBEL1254*, and *iLC915*. While strains

producing a recombinant protein were included in the simulations, production rates or yields were not evaluated. Nonetheless, the approach of Tomàs-Gamisans et al. (2016) resulted in the most comprehensive *P. pastoris* GEM yet. Similar to the ongoing *S. cerevisiae* work on a collaborative GEM refinement (Sánchez and Nielsen, 2015), *iMT1026* and the methodology that led to its creation should encourage the *P. pastoris* community to cooperate on improving the GEM for its chosen microorganism.

To date only one practical example of genome engineering based on COBRA assisted prediction from a *P. pastoris* GEM has been published. As mentioned previously, Nocon et al. (2014) used *PpaMBEL1254* by Sohn et al. (2010) as GEM and applied MOMA and FSEOF algorithms for prediction of suitable knock-out or overexpression targets, respectively. A total of nine predicted targets were assayed separately and five of them led to increased titers of the target recombinant protein. Interestingly, the employed GEM is based on DMSZ 70382 while the experiments were conducted with X-33, a reconstituted variant of the histidine-auxotroph GS115 (Ahmad et al., 2014). With DSMZ 70382 having been reclassified as *K. pastoris* and GS115 as *K. phaffii*, a certain evolutionary distance can be inferred (Kurtzman, 2009). Distinct differences on both the genomic and transcriptomic levels were recently detailed (Love et al., 2016). Potentially, this circumstance led to a lower accuracy of the GEM in this particular application. Nevertheless, the first implementation of GEM-based genetic engineering in *P. pastoris* was a success with product yields having been increased up to 40% (Nocon et al., 2014). It showcases the potential COBRA methods have for strain engineering in *P. pastoris*. The newly developed *iMT1026*, or future iterations of it, will enable more accurate target predictions and the combination of both *K. phaffii* and *K. pastoris* data should make it suitable for a wider range of applications.

7.3. Physiological insights

A key component to the improvement of a GEM is the inclusion of physiological data (Liu et al., 2010). Understanding the physiological implications and limitations of the host organism is essential for network construction in a COBRA approach (Schellenberger et al., 2011). For example, the capacity of the cellular membrane for anchoring membrane proteins has been incorporated into an *E. coli* GEM in order to increase its applicability for modeling expression of membrane bound proteins (J.K. Liu et al., 2014). Similarly, physiological insights have been used to refine *P. pastoris* GEM (Tomàs-Gamisans et al., 2016). Various genes involved in peroxisome biogenesis, membrane transport and central peroxisomal metabolic pathways were studied (Polupanov et al., 2011; Spong and Subramani, 1993; Yan et al., 2008) and its composition analyzed (Wriessnegger et al., 2007). Soon after *P. pastoris* suitability for recombinant protein production was established, related cellular physiological properties were investigated. Often, the goal was to understand factors limiting protein expression, folding and secretion (reviewed in Puxbaum et al. (2015)). As a result, a sizeable body of knowledge about peroxisomal and protein related physiological properties of *P. pastoris* is available. Here, we highlight some of the most recent achievements that further deepen these subjects or provide insights into as of yet less-studied physiological aspects.

Much progress has been made in understanding the transcriptional regulation of *AOX1* and other MUT genes (see Section 3.1). However, the exact mechanism of how the cell senses methanol remained unclear. Recently, two proteins homologs to the *WSC* (cell Wall integrity and Stress response Component) membrane protein family of *S. cerevisiae* (Verna et al., 1997) were found to mediate extracellular methanol sensing in *P. pastoris* (Ohsawa et al., 2017), as shown in Fig. 7(A). A total of three *WSC* genes were found on the *P. pastoris* genome, *PpWSC1*, *PpWSC2* and *PpWSC3*. Different deletion strains were created and their growth behavior as well as the transcriptional activity of methanol associated genes at different concentrations analyzed. In combination with complementation experiments, it was elucidated that *PpWSC1* and *PpWSC3*, but not *PpWSC2* are involved in methanol sensing. They showed different responses to changing methanol concentration, suggesting that *PpWSC1* is responsible for methanol sensing at low concentration (0.01 to 0.05% (w/w)) and *PpWSC3* for higher levels (0.1 to 0.5% (w/w)). Compared to the high range ethanol sensing mechanism of *S. cerevisiae* (Stanley et al., 2010) such a sensitive and two-tiered system is necessary due to the toxicity of intermediates of the methanol metabolism. *P. pastoris* must avoid the accumulation of toxic formaldehyde by fine-tuning the *AOX1* expression based on the methanol concentration. In the next step, knowledge about *WSC* involvement in cell wall integrity pathways (Levin, 2011) was used to identify a homolog of the signaling gene *ROM2* in *P. pastoris*. The homolog *PpROM2* was found and via a combination of knock-out and complementation studies, its involvement in methanol induced transcription and its binding to *PpWSC1/PpWSC2* was discovered. In consequence, a model was proposed by which *PpWCS1* and *PpWSC2* are transmembrane proteins that sense extracellular methanol and upon detection release *PpROM2* from their cytosolic tails. *PpROM2* translocates to the nucleus where it induces transcription of MUT genes. More than 30 mutated strains were constructed and protein-protein interactions assayed via protein engineering (Ohsawa et al., 2017). Thereby, important insights into the mechanism of methanol sensing were revealed. These insights open up a plethora of new genetic engineering opportunities. For instance, methanol sensing sensitivity could be altered to reduce the needed methanol concentration or the mechanism could be employed to express other non-methanol related pathways.

Although *P. pastoris* has an efficient native secretion apparatus, further optimization is possible. Puxbaum et al. (2016) investigated the secretion behavior of different recombinant proteins in regard to their cellular localization. It was found, that proteins transition from the cytosol to the medium predominately at the bud surface (Fig. 7(B)). Increased activity of the cell wall biogenesis and transport of involved proteins to the growing bud in *S. cerevisiae* has been reported before (Levin, 2011). However, so far no data was available regarding the cellular localization of exocytosis in budding yeasts. The secretory pathway of recombinant fluorescent fusion proteins was monitored in live cells via 4D fluorescence microscopy. During cultivation, cells were placed every few minutes on a confocal microscope and a z stack covering the entire cell was recorded. This approach allowed for the spatially and temporally resolved intracellular localization of the fusion proteins. As cultivation time increased, an increasing transition of fusion protein from the ER of the mother cell to the bud tip was visible. Interestingly, as soon as the ER in the bud was fully inherited from the mother cell, the fusion proteins were not transported to the bud surface anymore. While useful, the constant loss of signal due to secretion of fusion protein into the medium complicated this method. In the next step, recombinant and endogenously secreted proteins were fused to a membrane anchor domain to hold it in place. Target proteins were visualized via immunostaining. Again, 4D fluorescence microscopy was applied. The results clearly show, that both heterologous and native proteins were first secreted at the bud tip and later at the entire bud surface. It cannot be excluded that the procedures necessary for 4D fluorescence microscopy had an impact on cellular physiology, but the evidently high protein trafficking activity at the yeast bud (Levin, 2011) supports the findings of Puxbaum et al. (2016). Similarly, it has been found that in filamentous fungi exocytosis occurs predominantly at the growing hyphae tip (Shoji et al., 2014). Nevertheless, the present study elucidates the physiological properties of protein secretion in *P. pastoris*, and potentially all budding yeasts. It should therefore prove valuable for devising strategies to increase the secretion capacity by increasing the ratio of budding cells. These findings would also explain the improved secretion features of *P. pastoris* cells genetically engineered to exhibit prolonged G2 and M phases of cell division (Buchetics et al., 2011).

A three-step signal peptide was recently discovered (Heiss et al., 2015). Pursuing earlier studies on epx1 (Heiss et al., 2013), the

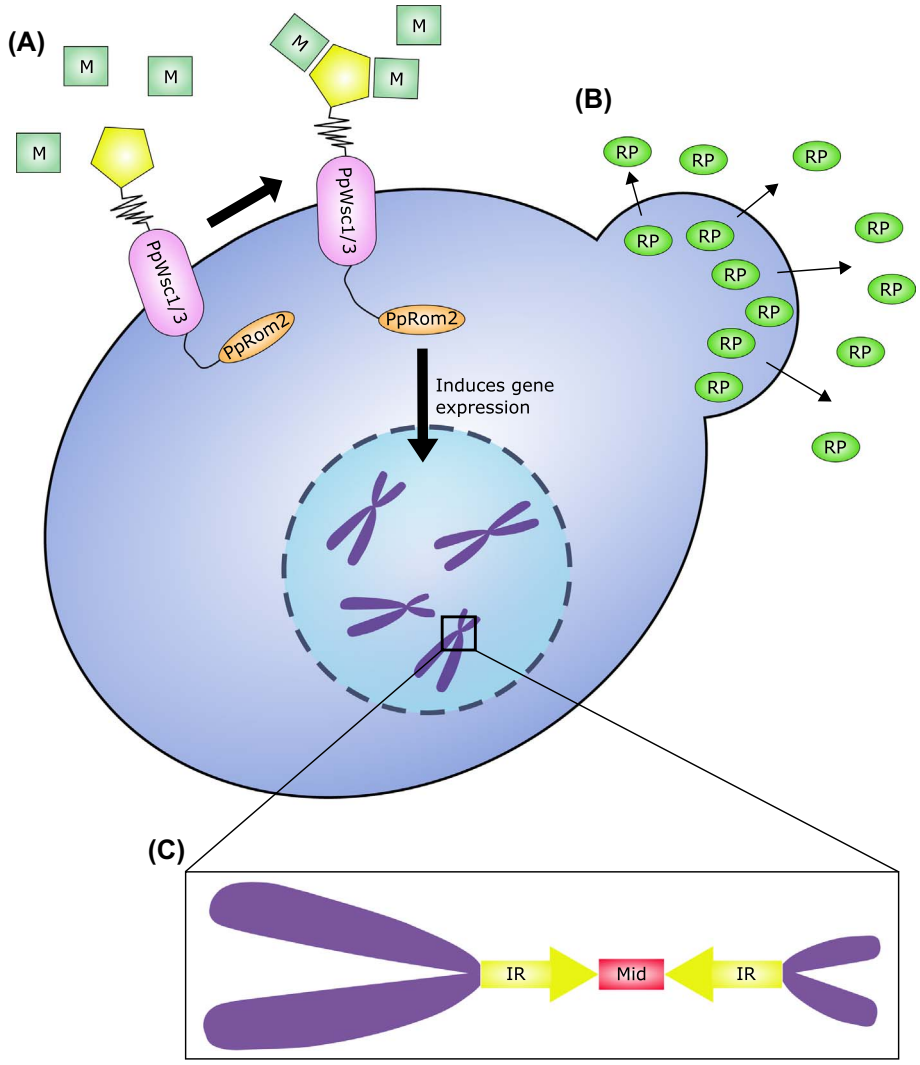

Fig. 7. Cartoon of three recent physiological studies. (A) Methanol sensing complex. *PpWsc1* and *PpWsc3* sense extracellular methanol. The attached *PpRom2* is suspected to translocate to the nucleus and activate methanol inducible gene expression. (B) Recombinant proteins (RP) are almost exclusively secreted at the bud tip. (C) The centromere structure was fully elucidated. It consists of a middle section (Mid) flanked by two inverted repeats (IR).

endogenous protein of highest abundance in the supernatant, its secretion signal was characterized. It was found that it contains a bipartite prosequence, requiring two endoprotease steps before the final signal sequence is generated. Interestingly, the cleavage site inside the bipartite prosequence could not be found in *EPX1* homologous of other yeasts. Modifications of the site lead to mistargeting of epx1 in some cases. Besides its physiological relevance, truncated versions of the epx1 secretion signal were also shown to facilitate efficient secretion of recombinant proteins. No detailed secretion rates were published so far, but initial results were promising and could lead to the development of an alternative to the *S. cerevisiae* α-mating factor secretion signal. In contrast, computational analysis and in vivo validation of various endogenous secretion signals yielded no improved secretion in comparison to the α-mating factor secretion signal (Massahi and Çalik, 2016). However, the epx1 secretion signal was not included in these studies.

Centromeres play an important role during cell division as the anchor point between cytoskeleton and chromosome. In particular the centromere of *S. cerevisiae* has been well characterized and used as model to better understand centromeres of higher eukaryotes like plants and humans (Blackburn and Szostak, 1984). Until recently, the centromere structure and position of *P. pastoris* was not fully elucidated. Applying chromatin immunoprecipitation (ChIP) sequencing, Coughlan et al. (2016) revealed the centromere organization (Fig. 7(C)). First, non-transcribed regions of the chromosomes were identified in available transcriptome data (Liang et al., 2012), reasoning that they should indicate the location of the centromeres. On each chromosome a 5–7 kb long non-transcribed region consisting of two ca. 2 kb inverted repeats (IR) separated by a ca. 1 kb non-repetitive middle section were found. Via a ChIP-Seq targeting the yeast centromeric histone 4, the proposed regions were confirmed to be the centromeres of *P. pastoris*. Their position on the chromosomes was in agreement with replication start data (Liachko et al., 2014). This setup differs markedly from the small point centromere *of S. cerevisiae* and is more similar to the organization found in *Candida tropicalis* and *Schizosaccharomyces pombe* (Coughlan et al., 2016). *P. pastoris* centromeres displayed only low similarity between chromosomes, making them more accessible to genetic manipulation. Centromeres can be applied for plasmid propagation in yeast and due to their low copy number are often used in conjunction with toxic targets (Rose et al., 1987). Engineering of synthetic chromosomes in *P. pastoris* will require further research, but the elucidation of the centromere organization is an important first step towards this application.

Compared to *S. cerevisiae*, exact information on biosynthetic pathways is less developed in non-conventional yeasts. To fill this knowledge gap, it is often assumed that *S. cerevisiae* pathways are conserved in other yeasts. In a computational approach, Förster et al. (2014) investigated the amino acid biosynthesis pathways of eight non-conventional yeasts, including *P. pastoris*. All analyzed yeasts had a more or less different enzymatic apparatus for amino acid biosynthesis, compared to *S. cerevisiae*. Isoenzymes were missing, different amount of gene copies were present and for some enzymes different intracellular localizations were predicted. The in silico approach was complemented by an experimental assessment of differences between *S. cerevisiae* and

P. pastoris L-leucine biosynthesis pathways. Thereby, two steps that occur predominantly in the mitochondrion of *S. cerevisiae* (Kohlhaw, 2003) could be located exclusively to the cytosol of *P. pastoris*. The results highlight the necessity to use inferred conservation of pathways from *S. cerevisiae* cautiously. Future *P. pastoris* GEM should implement these updated biosynthetic pathways.

8. Conclusion and future perspective

The key drawbacks of *P. pastoris*, clonal variability, lack of vectors and incomplete genome data, were significantly improved upon (Dikicioglu et al., 2014; Kelwick et al., 2014). Simultaneously, existing strengths like protein production were refined and as of yet untapped potential for metabolite synthesis revealed. The concerted scientific output since 2014 considerably accelerated the process of establishing *P. pastoris* as a microorganism suitable for systems metabolic engineering. In multiple studies the successful production of value-added metabolites from different classes was demonstrated. Concurrently, novel genetic engineering projects enabled new approaches for expression of glycoproteins, improved protein production characteristics and engineered traditional promoters to exhibit drastically new regulatory properties. The ability to fine-tune expression induction and strength has been distinctly augmented. Methanol-free *pAOX1* variants, repressible, fully synthetic and bidirectional promoters were introduced, exhibiting a wider range of expressional strengths. Clonal variability was systematically investigated and many causes for aberrant productivity and morphology were discovered. The elucidation of the underlying events and mechanisms in play will allow researchers to device strategies that prevent detrimental integration events from occurring. Thereby, the tedious screening procedures associated with *P. pastoris* might be significantly shortened. Popular genetic manipulation techniques like CRISPR and Golden Gate Assembly were successfully transferred, enabling the application of the newest synthetic biology methods. In addition, the novel concept of expressing pathways in a polycistronic way via 2A sequences could open up new avenues of metabolic engineering. The repertoire of available plasmids was markedly expanded with both integrative and episomal vectors. ARS based vectors displayed favourable productivity characteristics and a markedly reduced clonal variability. Applying NGS technology, the second generation of *P. pastoris* genome sequences was reported. Their high resolution, comprehensive functional annotation and unified nomenclature make them the ideal basis for future systems biology research. In particular transcriptomic and metabolomic analysis was used to gain further insights into cellular behavior under various environmental conditions. Also physiological discoveries were made, e.g. the first description of the methanol sensing apparatus. These physiological and biochemical results can be combined to form the next generation of *P. pastoris* GEM. Already, a renewed interest of refining existing GEM is discernible. Unfortunately, these efforts predated the publication of the second generation genome data.

Many promising optimization targets and potential products are available, but the compartmentalization of metabolite synthesis pathways to the peroxisome appears to be especially auspicious. The *P. pastoris* peroxisome has served as model peroxisome since many years and is therefore very well-characterized. In concert with the recent trend in yeast biotechnology to utilize peroxisomes for pathway compartmentalization (DeLoache et al., 2016; Shabbir Hussain et al., 2016), *P. pastoris* appears to be the ideal candidate to apply this strategy. As shown in Fig. 3, valuable metabolites like terpenoids and polyketides can be derived from the β-oxidation pathway in the peroxisome, and in some cases already have been. Especially terpenoids have gained much interest, due to the wide spectrum of applications they can be used for. By implementing tried model-based strategies from *S. cerevisiae*, terpenoid productivity in *P. pastoris* might be increased (Gruchattka et al., 2013; Gruchattka and Kayser, 2015). Already, it has been shown that for certain metabolites *P. pastoris* outperforms *S. cerevisiae* (Wriessnegger et al., 2014).

On the other hand, the well-established and characterized suitability of *P. pastoris* for expression of recombinant proteins can serve as the corner stone, upon which heterologous metabolic pathways can be implemented. Membrane proteins and glycoproteins that often are problematic in other heterologous host have been proven to be suited for high level active expression in *P. pastoris* (Byrne, 2015; Öberg et al., 2011). The large classes of CYP and CPR enzymes are often membrane anchored and are involved in a multitude of reactions that can be used for the production of valuable compounds. Published results indicate this to be a key factor in enabling higher yields in *P. pastoris* metabolic engineering for specific targets, compared to *E. coli* and *S. cerevisiae* (Chiang et al., 2016; Geier et al., 2012).

To fully realize systems metabolic engineering then, these discoveries need to be combined and expanded upon. For example, the new genomic, transcriptomic, metabolomic and physiological data should form the fundament upon which to create the newest generation of *P. pastoris* GEM. It has been shown that great potential lies in expanding GEM, facilitating more accurate predictions and the use of next-generation COBRA methods (King et al., 2015; Lee et al., 2012). The first reported model-based genetic engineering project in *P. pastoris* proved the potential this technology holds (Nocon et al., 2014). And while the updated version of www.pichiagenome.org has to be commended, there is still room for further developments. However, the inclusion of the new genome data will facilitate a stream lined transfer of knowledge between *P. pastoris* and *S. cerevisiae*. The *P. pastoris* community can now more easily benefit from the vast knowledge gathered on all systems biology levels of *S. cerevisiae*. This progress will need to go hand-in-hand with integrating new and improved genetic manipulation and synthetic biology techniques. Recent publications have provided the basis for rapid, modular and in vivo vector assembly, as well as shown more accurate ways of manipulating the genome. Via genome-sequencing or genotyping a large library of clones, further insights into random integrations might be obtained (Kegel et al., 2006; Zhang et al., 2014).

Groundbreaking studies, like point-of-care drug synthesis in a microfluidic reactor or detection of single-cell protein efflux, demonstrate the growing interest of scientists in this non-conventional yeast (Landry et al., 2017; Perez-Pinera et al., 2016). It therefore appears that the future of *P. pastoris* is bright and full of possibilities, but it will require the concerted effort of its scientific community to realize this potential.

Acknowledgments

Grants from the Federal State of North Rhine-Westphalia for the CLIB-Graduate Cluster Industrial Biotechnology are gratefully acknowledged.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 685778.

This work has been supported by the Federal Ministry of Science, Research and Economy (BMWFW), the Federal Ministry of Traffic, Innovation and Technology (BMVIT), the Styrian Business Promotion Agency SFG, the Standortagentur Tirol, the Government of Lower Austria and Business Agency Vienna through the COMET-Funding Program managed by the Austrian Research Promotion Agency FFG.

References

Abad, S., Kitz, K., Hörmann, A., Schreiner, U., Hartner, F.S., Glieder, A., 2010a. Real-time PCR-based determination of gene copy numbers in *Pichia pastoris*. Biotechnol. J. 5, 413–420. http://dx.doi.org/10.1002/biot.200900233.

Abad, S., Nahalka, J., Bergler, G., Arnold, S.A., Speight, R., Fotheringham, I., Nidetzky, B., Glieder, A., 2010b. Stepwise engineering of a *Pichia pastoris* D-amino acid oxidase whole cell catalyst. Microb. Cell Factories 9, 24. http://dx.doi.org/10.1186/1475-2859-9-24.

Ahmad, M., Hirz, M., Pichler, H., Schwab, H., 2014. Protein expression in *Pichia pastoris*:

recent achievements and perspectives for heterologous protein production. Appl. Microbiol. Biotechnol. 98, 5301–5317. http://dx.doi.org/10.1007/s00253-014-5732-5.

Antonenkov, V.D., Hiltunen, J.K., 2012. Transfer of metabolites across the peroxisomal membrane. Biochim. Biophys. Acta Mol. basis Dis. 1822, 1374–1386. http://dx.doi.org/10.1016/j.bbadis.2011.12.011.

Araya-Garay, J.M., Ageitos, J.M., Vallejo, J.A., Veiga-Crespo, P., Sánchez-Pérez, A., Villa, T.G., 2012a. Construction of a novel *Pichia pastoris* strain for production of xanthophylls. AMB Express 2, 24. http://dx.doi.org/10.1186/2191-0855-2-24.

Araya-Garay, J.M., Feijoo-Siota, L., Rosa-Dos-Santos, F., Veiga-Crespo, P., Villa, T.G., 2012b. Construction of new *Pichia pastoris* X-33 strains for production of lycopene and β-carotene. Appl. Microbiol. Biotechnol. 93, 2483–2492. http://dx.doi.org/10.1007/s00253-011-3764-7.

Asadollahi, M.A., Maury, J., Schalk, M., Clark, A., Nielsen, J., 2010. Enhancement of farnesyl diphosphate pool as direct precursor of sesquiterpenes through metabolic engineering of the mevalonate pathway in *Saccharomyces cerevisiae*. Biotechnol. Bioeng. 106, 86–96. http://dx.doi.org/10.1002/bit.22668.

Avalos, J.L., Fink, G.R., Stephanopoulos, G., 2013. Compartmentalization of metabolic pathways in yeast mitochondria improves the production of branched-chain alcohols. Nat. Biotechnol. 31, 335–341. http://dx.doi.org/10.1038/nbt.2509.

Aw, R., Polizzi, K.M., 2013. Can too many copies spoil the broth? Microb. Cell Factories 12, 128. http://dx.doi.org/10.1186/1475-2859-12-128.

Aw, R., Polizzi, K.M., 2016. Liquid PTVA: a faster and cheaper alternative for generating multi-copy clones in *Pichia pastoris*. Microb. Cell Factories 15, 29. http://dx.doi.org/10.1186/s12934-016-0432-8.

Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K.A., Tomita, M., Wanner, B.L., Mori, H., 2006. Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. Mol. Syst. Biol. 2, 2006.0008. http://dx.doi.org/10.1038/msb4100050.

Barnes, S., 2010. The biochemistry, chemistry and physiology of the isoflavones in soybeans and their food products. Lymphat. Res. Biol. 8, 89–98. http://dx.doi.org/10.1089/lrb.2009.0030.

Baumann, K., Carnicer, M., Dragosits, M., Graf, A.B., Stadlmann, J., Jouhten, P., Maaheimo, H., Gasser, B., Albiol, J., Mattanovich, D., Ferrer, P., 2010. A multi-level study of recombinant *Pichia pastoris* in different oxygen conditions. BMC Syst. Biol. 4, 141. http://dx.doi.org/10.1186/1752-0509-4-141.

Baumann, K., Adelantado, N., Lang, C., Mattanovich, D., Ferrer, P., 2011. Protein trafficking, ergosterol biosynthesis and membrane physics impact recombinant protein secretion in *Pichia pastoris*. Microb. Cell Factories 10, 93. http://dx.doi.org/10.1186/1475-2859-10-93.

Bayer, T.S., Widmaier, D.M., Temme, K., Mirsky, E.A., Santi, D.V., Voigt, C.A., 2009. Synthesis of methyl halides from biomass using engineered microbes. J. Am. Chem. Soc. 131, 6508–6515. http://dx.doi.org/10.1021/ja809461u.

Belhaj, K., Chaparro-Garcia, A., Kamoun, S., Patron, N.J., Nekrasov, V., 2015. Editing plant genomes with CRISPR/Cas9. Curr. Opin. Biotechnol. 32, 76–84. http://dx.doi.org/10.1016/j.copbio.2014.11.007.

Bhataya, A., Schmidt-Dannert, C., Lee, P.C., 2009. Metabolic engineering of *Pichia pastoris* X-33 for lycopene production. Process Biochem. 44, 1095–1102. http://dx.doi.org/10.1016/j.procbio.2009.05.012.

Bill, R.M., 2014. Playing catch-up with *Escherichia coli*: using yeast to increase success rates in recombinant protein production experiments. Front. Microbiol. 5, 85. http://dx.doi.org/10.3389/fmicb.2014.00085.

Blackburn, E.H., Szostak, J.W., 1984. The molecular structure of centromeres and telomeres. Annu. Rev. Biochem. 53, 163–194. http://dx.doi.org/10.1146/annurev.bi.53.070184.001115.

Bobrowicz, P., Davidson, R.C., Li, H., Potgieter, T.I., Nett, J.H., Hamilton, S.R., Stadheim, T.A., Miele, R.G., Bobrowicz, B., Mitchell, T., Rausch, S., Renfer, E., Wildt, S., 2004. Engineering of an artificial glycosylation pathway blocked in core oligosaccharide assembly in the yeast *Pichia pastoris*: production of complex humanized glycoproteins with terminal galactose. Glycobiology 14, 757–766. http://dx.doi.org/10.1093/glycob/cwh104.

Bruggeman, F.J., Westerhoff, H.V., 2007. The nature of systems biology. Trends Microbiol. 15, 45–50. http://dx.doi.org/10.1016/j.tim.2006.11.003.

Buchetics, M., Dragosits, M., Maurer, M., Rebnegger, C., Porro, D., Sauer, M., Gasser, B., Mattanovich, D., 2011. Reverse engineering of protein secretion by uncoupling of cell cycle phases from growth. Biotechnol. Bioeng. 108, 2403–2412. http://dx.doi.org/10.1002/bit.23198.

Byrne, B., 2015. *Pichia pastoris* as an expression host for membrane protein structural biology. Curr. Opin. Struct. Biol. 32, 9–17. http://dx.doi.org/10.1016/j.sbi.2015.01.005.

Cámara, E., Albiol, J., Ferrer, P., 2016. Droplet digital PCR-aided screening and characterization of *Pichia pastoris* multiple gene copy strains. Biotechnol. Bioeng. 113, 1542–1551. http://dx.doi.org/10.1002/bit.25916.

Camattari, A., Goh, A., Yip, L.Y., Tan, A.H.M., Ng, S.W., Tran, A., Liu, G., Liachko, I., Dunham, M.J., Rancati, G., 2016. Characterization of a panARS-based episomal vector in the methylotrophic yeast *Pichia pastoris* for recombinant protein production and synthetic biology applications. Microb. Cell Factories 15, 139. http://dx.doi.org/10.1186/s12934-016-0540-5.

Cankar, K., Van Houwelingen, A., Goedbloed, M., Renirie, R., De Jong, R.M., Bouwmeester, H., Bosch, D., Sonke, T., Beekwilder, J., 2014. Valencene oxidase CYP706M1 from Alaska cedar (*Callitropsis nootkatensis*). FEBS Lett. 588, 1001–1007. http://dx.doi.org/10.1016/j.febslet.2014.01.061.

Carnicer, M., Ten Pierick, A., Van Dam, J., Heijnen, J.J., Albiol, J., Van Gulik, W., Ferrer, P., 2012. Quantitative metabolomics analysis of amino acid metabolism in recombinant *Pichia pastoris* under different oxygen availability conditions. Microb. Cell Factories 11, 1. http://dx.doi.org/10.1186/1475-2859-11-83.

Casini, A., Storch, M., Baldwin, G.S., Ellis, T., 2015. Bricks and blueprints: methods and standards for DNA assembly. Nat. Rev. Mol. Cell Biol. 16, 568–576. http://dx.doi.org/10.1038/nrm4014.

Caspeta, L., Shoaie, S., Agren, R., Nookaew, I., Nielsen, J., 2012. Genome-scale metabolic reconstructions of Pichia stipitis and *Pichia pastoris* and in silico evaluation of their potentials. BMC Syst. Biol. 6, 24. http://dx.doi.org/10.1186/1752-0509-6-24.

Cereghino, J.L., Cregg, J.M., 2000. Heterologous protein expression in the methylotrophic yeast *Pichia pastoris*. FEMS Microbiol. Rev. 24, 45–66. http://dx.doi.org/10.1111/j.1574-6976.2000.tb00532.x.

Chang, T.S., Chao, S.Y., Chen, Y.C., 2013. Production of ortho-hydroxydaidzein derivatives by a recombinant strain of *Pichia pastoris* harboring a cytochrome P450 fusion gene. Process Biochem. 48, 426–429. http://dx.doi.org/10.1016/j.procbio.2013.02.014.

Chen, W., Smeekens, J.M., Wu, R., 2014. Comprehensive analysis of protein *N*-glycosylation sites by combining chemical deglycosylation with LC − MS. J. Proteome Res. 13, 1466–1473. http://dx.doi.org/10.1021/pr401000c.

Cheng, H., Lv, J., Wang, H., Wang, B., Li, Z., Deng, Z., 2014. Genetically engineered *Pichia pastoris* yeast for conversion of glucose to xylitol by a single-fermentation process. Appl. Microbiol. Biotechnol. 98, 3539–3552. http://dx.doi.org/10.1007/s00253-013-5501-x.

Chiang, C.M., Ding, H.Y., Lu, J.Y., Chang, T.S., 2016. Biotransformation of isoflavones daidzein and genistein by recombinant *Pichia pastoris* expressing membrane-anchoring and reductase fusion chimeric CYP105D7. J. Taiwan Inst. Chem. Eng. 60, 26–31. http://dx.doi.org/10.1016/j.jtice.2015.10.015.

Choi, B., Bobrowicz, P., Davidson, R.C., Hamilton, S.R., Kung, D.H., Li, H., Miele, R.G., Nett, J.H., Wildt, S., Tillman, U., 2003. Use of combinatorial genetic libraries to humanize N-linked glycosylation in the yeast *Pichia pastoris*. Proc. Natl. Acad. Sci. U. S. A. 100, 5022–5027. http://dx.doi.org/10.1073/pnas.0931263100.

Choi, H.S., Lee, S.Y., Kim, T.Y., Woo, H.M., 2010. In silico identification of gene amplification targets for improvement of lycopene production. Appl. Environ. Microbiol. 76, 3097–3105. http://dx.doi.org/10.1128/AEM.00115-10.

Choi, S.Y., Yoon, K.H., Lee, J. Il, Mitchell, R.J., 2015. Violacein: properties and production of a versatile bacterial pigment. Biomed. Res. Int. 2015. http://dx.doi.org/10.1155/2015/465056.

Chung, B.K.S., Selvarasu, S., Camattari, A., Ryu, J., Lee, H., Ahn, J., Lee, H., Lee, D., 2010. Genome-scale metabolic reconstruction and in silico analysis of methylotrophic yeast *Pichia pastoris* for strain improvement. Microb. Cell Factories 9.

Ciofalo, V., Barton, N., Kreps, J., Coats, I., Shanahan, D., 2006. Safety evaluation of a lipase enzyme preparation, expressed in *Pichia pastoris*, intended for use in the degumming of edible vegetable oil. Regul. Toxicol. Pharmacol. 45, 1–8. http://dx.doi.org/10.1016/j.yrtph.2006.02.001.

Claes, K., Vandewalle, K., Laukens, B., Laeremans, T., Vosters, O., Langer, I., Parmentier, M., Steyaert, J., Callewaert, N., 2016. Modular integrated secretory system engineering in *Pichia pastoris* to enhance G-protein coupled receptor expression. ACS Synth. Biol. 5, 1070–1075. http://dx.doi.org/10.1021/acssynbio.6b00032.

Clare, J., Rayment, F., Sreekrishna, K., Romanos, M.A., 1991. High-level expression of tetanus toxin fragment C in *Pichia pastoris* strains containing multiple tandem integrations of the gene. Nat. Biotechnol. 9, 455–460. http://dx.doi.org/10.1038/nbt0591-455.

Clasquin, M.F., Melamud, E., Singer, A., Gooding, J.R., Xu, X., Dong, A., Cui, H., Campagna, S.R., Savchenko, A., Yakunin, A.F., Rabinowitz, J.D., Caudy, A.A., 2011. Riboneogenesis in yeast. Cell 145, 969–980. http://dx.doi.org/10.1016/j.cell.2011.05.022.

Corchero, J.L., Gasser, B., Resina, D., Smith, W., Parrilli, E., Vázquez, F., Abasolo, I., Giuliani, M., Jäntti, J., Ferrer, P., Saloheimo, M., Mattanovich, D., Schwartz, S., Tutino, M.L., Villaverde, A., 2013. Unconventional microbial systems for the cost-efficient production of high-quality protein therapeutics. Biotechnol. Adv. 31, 140–153. http://dx.doi.org/10.1016/j.biotechadv.2012.09.001.

Couderc, R., Baratti, J., 1980. Oxidation of methanol by the yeast, *Pichia pastoris*: purification and properties of the alcohol oxidase. Agric. Biol. Chem. 44, 2279–2289.

Coughlan, A.Y., Hanson, S.J., Byrne, K.P., Wolfe, K.H., 2016. Centromeres of the yeast *Komagataella phaffii* (*Pichia pastoris*) have a simple inverted-repeat structure. Genome Biol. Evol. 8, 2482–2492. http://dx.doi.org/10.1093/gbe/evw178.

Crampton, N., Bonass, W.A., Kirkham, J., Rivetti, C., Thomson, N.H., 2006. Collision events between RNA polymerases in convergent transcription studied by atomic force microscopy. Nucleic Acids Res. 34, 5416–5425. http://dx.doi.org/10.1093/nar/gkl668.

Cregg, J., Barringer, K., Hessler, A., Madden, K., 1985. *Pichia pastoris* as a host system for transformations. Mol. Cell. Biol. 5, 3376–3385. http://dx.doi.org/10.1128/MCB.5.12.3376.

Curran, K.A., Karim, A.S., Gupta, A., Alper, H.S., 2013. Use of expression-enhancing terminators in *Saccharomyces cerevisiae* to increase mRNA half-life and improve gene expression control for metabolic engineering applications. Metab. Eng. 19, 88–97. http://dx.doi.org/10.1016/j.ymben.2013.07.001.

Dai, Z., Liu, Y., Zhang, X., Shi, M., Wang, B., Wang, D., Huang, L., Zhang, X., 2013. Metabolic engineering of *Saccharomyces cerevisiae* for production of ginsenosides. Metab. Eng. 20, 146–156. http://dx.doi.org/10.1016/j.ymben.2013.10.004.

Daley, J.M., Palmbos, P.L., Wu, D., Wilson, T.E., 2005. Nonhomologous end joining in yeast. Annu. Rev. Genet. 39, 431–451. http://dx.doi.org/10.1146/annurev.genet.39.073003.113340.

Davidson, R.C., Nett, J.H., Renfer, E., Li, H., Stadheim, T.A., Miller, B.J., Miele, R.G., Hamilton, S.R., Choi, B.K., Mitchell, T.I., Wildt, S., 2004. Functional analysis of the ALG3 gene encoding the Dol-P-Man: Man5GlcNAc2-PP-Dol mannosyltransferase enzyme of *P. pastoris*. Glycobiology 14, 399–407. http://dx.doi.org/10.1093/glycob/cwh023.

Day, C.D., Lee, E., Kobayashi, J., Holappa, L.D., Albert, H., Ow, D.W., 2000. Transgene

integration into the same chromosome location can produce alleles that express at a predictable level, or alleles that are differentially silenced. Genes Dev. 14, 2869–2880. http://dx.doi.org/10.1101/gad.849600.
De Felipe, P., Luke, G.A., Hughes, L.E., Gani, D., Halpin, C., Ryan, M.D., 2006. *E unum pluribus*: multiple proteins from a self-processing polyprotein. Trends Biotechnol. 24, 68–75. http://dx.doi.org/10.1016/j.tibtech.2005.12.006.
de Lima, P.B.A., Mulder, K.C.L., Melo, N.T.M., Carvalho, L.S., Menino, G.S., Mulinari, E., de Castro, V.H., Dos Reis, T.F., Goldman, G.H., Magalhães, B.S., Parachin, N.S., 2016. Novel homologous lactate transporter improves L-lactic acid production from glycerol in recombinant strains of *Pichia pastoris*. Microb. Cell Factories 15, 158. http://dx.doi.org/10.1186/s12934-016-0557-9.
De Schutter, K., Lin, Y.-C., Tiels, P., Van Hecke, A., Glinka, S., Weber-Lehmann, J., Rouzé, P., Van de Peer, Y., Callewaert, N., 2009. Genome sequence of the recombinant protein production host *Pichia pastoris*. Nat. Biotechnol. 27, 561–566. http://dx.doi.org/10.1038/nbt.1544.
Dean, N., 1999. Asparagine-linked glycosylation in the yeast Golgi. Biochim. Biophys. Acta, Gen. Subj. 1426, 309–322. http://dx.doi.org/10.1016/S0304-4165(98)00132-9.
Delic, M., Rebnegger, C., Wanka, F., Puxbaum, V., Haberhauer-Troyer, C., Hann, S., Köllensperger, G., Mattanovich, D., Gasser, B., 2012. Oxidative protein folding and unfolded protein response elicit differing redox regulation in endoplasmic reticulum and cytosol of yeast. Free Radic. Biol. Med. 52, 2000–2012. http://dx.doi.org/10.1016/j.freeradbiomed.2012.02.048.
Delic, M., Mattanovich, D., Gasser, B., 2013. Repressible promoters - a novel tool to generate conditional mutants in *Pichia pastoris*. Microb. Cell Factories 12, 6. http://dx.doi.org/10.1186/1475-2859-12-6.
DeLoache, W.C., Russ, Z.N., Dueber, J.E., 2016. Towards repurposing the yeast peroxisome for compartmentalizing heterologous metabolic pathways. Nat. Commun. 7, 11152. http://dx.doi.org/10.1038/ncomms11152.
DiCarlo, J.E., Norville, J.E., Mali, P., Rios, X., Aach, J., Church, G.M., 2013. Genome engineering in *Saccharomyces cerevisiae* using CRISPR-Cas systems. Nucleic Acids Res. 41, 4336–4343. http://dx.doi.org/10.1093/nar/gkt135.
Dikicioglu, D., Wood, V., Rutherford, K.M., McDowall, M.D., Oliver, S.G., 2014. Improving functional annotation for industrial microbes: a case study with *Pichia pastoris*. Trends Biotechnol. 1–4. http://dx.doi.org/10.1016/j.tibtech.2014.05.003.
Dixon, R.A., Pasinetti, G.M., 2010. Flavonoids and isoflavonoids: from plant biology to agriculture and neuroscience. Plant Physiol. 154, 453–457. http://dx.doi.org/10.1104/pp.110.161430.
Donsbach, K., Rück-Braun, K., 2005. Combinatorial biosynthesis of polyketides. Nat. Rev. Microbiol. 3, 925–936. http://dx.doi.org/10.1002/9783527619979.ch46.
Douin, V., Bornes, S., Creancier, L., Rochaix, P., Favre, G., Prats, A.-C., Couderc, B., 2004. Use and comparison of different internal ribosomal entry sites (IRES) in tricistronic retroviral vectors. BMC Biotechnol. 4, 16. http://dx.doi.org/10.1186/1472-6750-4-16.
Dragosits, M., Stadlmann, J., Albiol, J., Baumann, K., Maurer, M., Gasser, B., Sauer, M., Altmann, F., Ferrer, P., Mattanovich, D., 2009. The effect of temperature on the proteome of recombinant *Pichia pastoris*. J. Proteome Res. 8, 1380–1392. http://dx.doi.org/10.1021/pr8007623.
Dragosits, M., Stadlmann, J., Graf, A., Gasser, B., Maurer, M., Sauer, M., Kreil, D.P., Altmann, F., Mattanovich, D., 2010. The response to unfolded protein is involved in osmotolerance of *Pichia pastoris*. BMC Genomics 11, 207. http://dx.doi.org/10.1186/1471-2164-11-207.
Du, M., Battles, M.B., Nett, J.H., 2012. A color-based stable multi-copy integrant selection system for *Pichia pastoris* using the attenuated ADE1 and ADE2 genes as auxotrophic markers. Bioeng. Bugs 3, 32–37. http://dx.doi.org/10.4161/bbug.3.1.17936.
Edwards-Jones, B., Aw, R., Barton, G.R., Tredwell, G.D., Bundy, J.G., Leak, D.J., 2015. Translational arrest due to cytoplasmic redox stress delays adaptation to growth on methanol and heterologous protein expression in a typical fed-batch culture of *Pichia pastoris*. PLoS One 10, e0119637. http://dx.doi.org/10.1371/journal.pone.0119637.
Eisenberg, A., Seip, J.E., Gavagan, J.E., Payne, M.S., Anton, D.L., Dicosimo, R., 1997. Pyruvic acid production using methylotrophic as catalyst yeast transformants. J. Mol. Catal. 2, 223–232. http://dx.doi.org/10.1016/S1381-1177(96)00021-5.
Emmerstorfer, A., Wimmer-Teubenbacher, M., Wriessnegger, T., Leitner, E., Müller, M., Kaluzna, I., Schürmann, M., Mink, D., Zellnig, G., Schwab, H., Pichler, H., 2015. Over-expression of ICE2 stabilizes cytochrome P450 reductase in *Saccharomyces cerevisiae* and *Pichia pastoris*. Biotechnol. J. 10, 623–635. http://dx.doi.org/10.1002/biot.201400780.
Essani, K., Glieder, A., Geier, M., 2015. Combinatorial pathway assembly in yeast. AIMS Bioeng. 2, 423–436. http://dx.doi.org/10.3934/bioeng.2015.4.423.
Feng, P., Shang, Y., Cen, K., Wang, C., 2015. Fungal biosynthesis of the bibenzoquinone oosporein to evade insect immunity. Proc. Natl. Acad. Sci. U. S. A. 112, 11365–11370. http://dx.doi.org/10.1073/pnas.1503200112.
Fernandes, P., Cruz, A., Angelova, B., Pinheiro, H.M., Cabral, J.M.S., 2003. Microbial conversion of steroid compounds: recent developments. Enzym. Microb. Technol. 32, 688–705. http://dx.doi.org/10.1016/S0141-0229(03)00029-2.
Fliss, A.E., Benzeno, S., Rao, J., Caplan, A.J., 2000. Control of estrogen receptor ligand binding by Hsp90. J. Steroid Biochem. Mol. Biol. 72, 223–230. http://dx.doi.org/10.1016/S0960-0760(00)00037-6.
Fogg, C., Lustig, S., Whitbeck, J.C., Eisenberg, R.J., Cohen, G.H., Moss, B., 2004. Protective immunity to vaccinia virus induced by vaccination with multiple recombinant outer membrane proteins of intracellular and extracellular virions. J. Virol. 78, 10230–10237. http://dx.doi.org/10.1128/JVI.78.19.10230.
Förster, J., Halbfeld, C., Zimmermann, M., Blank, L.M., 2014. A blueprint of the amino acid biosynthesis network of hemiascomycetes. FEMS Yeast Res. 14, 1090–1100. http://dx.doi.org/10.1111/1567-1364.12205.
Fraatz, M.A., Berger, R.G., Zorn, H., 2009. Nootkatone-a biotechnological challenge. Appl. Microbiol. Biotechnol. 83, 35–41. http://dx.doi.org/10.1007/s00253-009-1968-x.
Galli, A., Schiestl, R.H., 1996. Hydroxyurea induces recombination in dividing but not in G1 or G2 cell cycle arrested yeast cells. Mutat. Res. 354, 69–75. http://dx.doi.org/10.1016/0027-5107(96)00037-1.
Gao, L., Cai, M., Shen, W., Xiao, S., Zhou, X., Zhang, Y., 2013. Engineered fungal polyketide biosynthesis in *Pichia pastoris*: a potential excellent host for polyketide production. Microb. Cell Factories 12, 77. http://dx.doi.org/10.1186/1475-2859-12-77.
Gasser, B., Maurer, M., Gach, J., Kunert, R., Mattanovicha, D., 2006. Engineering of *Pichia pastoris* for improved production of antibody fragments. Biotechnol. Bioeng. 94, 353–361. http://dx.doi.org/10.1002/bit.
Gasser, B., Maurer, M., Rautio, J., Sauer, M., Bhattacharyya, A., Saloheimo, M., Penttilä, M., Mattanovich, D., 2007. Monitoring of transcriptional regulation in *Pichia pastoris* under protein production conditions. BMC Genomics 8, 179. http://dx.doi.org/10.1186/1471-2164-8-179.
Gasser, B., Dragosits, M., Mattanovich, D., 2010. Engineering of biotin-prototrophy in *Pichia pastoris* for robust production processes. Metab. Eng. 12, 573–580. http://dx.doi.org/10.1016/j.ymben.2010.07.002.
Gavagan, J.E., Fager, S.K., Seip, J.E., Clark, D.S., Payne, M.S., Anton, D.L., Dicosimo, R., 1997. Chemoenzymic synthesis of *N*-(phosphonomethyl) glycine. J. Organomet. Chem. 62, 5419–5427. http://dx.doi.org/10.1021/jo970455f.
Gavira, C., Höfer, R., Lesot, A., Lambert, F., Zucca, J., Werck-Reichhart, D., 2013. Challenges and pitfalls of P450-dependent (+)-valencene bioconversion by *Saccharomyces cerevisiae*. Metab. Eng. 18, 25–35. http://dx.doi.org/10.1016/j.ymben.2013.02.003.
Geier, M., Braun, A., Emmerstorfer, A., Pichler, H., Glieder, A., 2012. Production of human cytochrome P450 2D6 drug metabolites with recombinant microbes - a comparative study. Biotechnol. J. 7, 1346–1358. http://dx.doi.org/10.1002/biot.201200187.
Geier, M., Brandner, C., Strohmeier, G.A., Hall, M., Hartner, F.S., Glieder, A., 2015a. Engineering *Pichia pastoris* for improved NADH regeneration: a novel chassis strain for whole-cell catalysis. Beilstein J. Org. Chem. 11, 1741–1748. http://dx.doi.org/10.3762/bjoc.11.190.
Geier, M., Fauland, P.C., Vogl, T., Glieder, A., 2015b. Compact multi enzyme pathways in *Pichia pastoris*. Chem. Commun. 51, 1643–1646. http://dx.doi.org/10.1039/C4CC08502G.
Generoso, W.C., Gottardi, M., Oreb, M., Boles, E., 2016. Simplified CRISPR-Cas genome editing for *Saccharomyces cerevisiae*. J. Microbiol. Methods 127, 203–205. http://dx.doi.org/10.1016/j.mimet.2016.06.020.
Gleeson, M.A., Sudbery, P.E., 1988. The methylotrophic yeasts. Yeast 4, 1–15. http://dx.doi.org/10.1002/yea.320040102.
Graf, A., Gasser, B., Dragosits, M., Sauer, M., Leparc, G.G., Tüchler, T., Kreil, D.P., Mattanovich, D., 2008. Novel insights into the unfolded protein response using *Pichia pastoris* specific DNA microarrays. BMC Genomics 9, 390. http://dx.doi.org/10.1186/1471-2164-9-390.
Gruchattka, E., Kayser, O., 2015. In vivo validation of in silico predicted metabolic engineering strategies in yeast: disruption of alpha-ketoglutarate dehydrogenase and expression of ATP-citrate lyase for terpenoid production. PLoS One 10, 1–25. http://dx.doi.org/10.1371/journal.pone.0144981.
Gruchattka, E., Hädicke, O., Klamt, S., Schütz, V., Kayser, O., 2013. In silico profiling of *Escherichia coli* and *Saccharomyces cerevisiae* as terpenoid factories. Microb. Cell Factories 12, 84. http://dx.doi.org/10.1186/1475-2859-12-84.
Guerfal, M., Ryckaert, S., Jacobs, P.P., Ameloot, P., Van Craenenbroeck, K., Derycke, R., Callewaert, N., 2010. The HAC1 gene from *Pichia pastoris*: characterization and effect of its overexpression on the production of secreted, surface displayed and membrane proteins. Microb. Cell Factories 9, 49. http://dx.doi.org/10.1186/1475-2859-9-49.
Guirouilh-Barbat, J., Huck, S., Bertrand, P., Pirzio, L., Desmaze, C., Sabatier, L., Lopez, B.S., 2004. Impact of the KU80 pathway on NHEJ-induced genome rearrangements in mammalian cells. Mol. Cell 14, 611–623. http://dx.doi.org/10.1016/j.molcel.2004.05.008.
Hamilton, S.R., Gerngross, T.U., 2007. Glycosylation engineering in yeast: the advent of fully humanized yeast. Curr. Opin. Biotechnol. 18, 387–392. http://dx.doi.org/10.1016/j.copbio.2007.09.001.
Hamilton, S.R., Davidson, R.C., Sethuraman, N., Nett, J.H., Jiang, Y., Rios, S., Bobrowicz, P., Stadheim, T.A., Li, H., Choi, B., Hopkins, D., Wischnewski, H., Roser, J., Mitchell, T., Strawbridge, R.R., Hoopes, J., Wildt, S., Gerngross, T.U., 2006. Humanization of yeast to produce complex terminally sialylated glycoproteins. Science 313, 1441–1443. http://dx.doi.org/10.1126/science.1130256.
Hartner, F.S., Ruth, C., Langenegger, D., Johnson, S.N., Hyka, P., Lin-Cereghino, G.P., Lin-Cereghino, J., Kovar, K., Cregg, J.M., Glieder, A., 2008. Promoter library designed for fine-tuned gene expression in *Pichia pastoris*. Nucleic Acids Res. 36, e76. http://dx.doi.org/10.1093/nar/gkn369.
He, Y., Cox, R.J., 2016. The molecular steps of citrinin biosynthesis in fungi. Chem. Sci. 7, 2119–2127. http://dx.doi.org/10.1039/C5SC04027B.
Heiss, S., Maurer, M., Hahn, R., Mattanovich, D., Gasser, B., 2013. Identification and deletion of the major secreted protein of *Pichia pastoris*. Appl. Microbiol. Biotechnol. 97, 1241–1249. http://dx.doi.org/10.1007/s00253-012-4260-4.
Heiss, S., Puxbaum, V., Gruber, C., Altmann, F., Mattanovich, D., Gasser, B., 2015. Multistep processing of the secretion leader of the extracellular protein Epx1 in *Pichia pastoris* and implications for protein localization. Microbiology 161, 1356–1368. http://dx.doi.org/10.1099/mic.0.000105.
Hemmerich, J., Adelantado, N., Barrigón, J.M., Ponte, X., Hörmann, A., Ferrer, P., Kensy, F., Valero, F., 2014. Comprehensive clone screening and evaluation of fed-batch strategies in a microbioreactor and lab scale stirred tank bioreactor system: application on *Pichia pastoris* producing *Rhizopus oryzae* lipase. Microb. Cell Factories 13, 36. http://dx.doi.org/10.1186/1475-2859-13-36.

Hesketh, A.R., Castrillo, J.I., Sawyer, T., Archer, D.B., Oliver, S.G., 2013. Investigating the physiological response of *Pichia* (*Komagataella*) *pastoris* GS115 to the heterologous expression of misfolded proteins using chemostat cultures. Appl. Microbiol. Biotechnol. 97, 9747–9762. http://dx.doi.org/10.1007/s00253-013-5186-1.

Heyland, J., Fu, J., Blank, L.M., Schmid, A., 2011. Carbon metabolism limits recombinant protein production in *Pichia pastoris*. Biotechnol. Bioeng. 108, 1942–1953. http://dx.doi.org/10.1002/bit.23114.

Hori, K., Okano, S., Sato, F., 2016. Efficient microbial production of stylopine using a *Pichia pastoris* expression system. Sci Rep 6, 22201. http://dx.doi.org/10.1038/srep22201.

Inan, M., Meagher, M.M., 2001. Non-repressing carbon sources for alcohol oxidase (AOX1) promoter of *Pichia pastoris*. J. Biosci. Bioeng. 92, 585–589.

Ingolia, N.T., Ghaemmaghami, S., Newman, J.R.S., Weissman, J.S., 2009. Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. Science 324, 218–223. http://dx.doi.org/10.1126/science.1168978.

Invitrogen, 2010. Pichia Fermentation Process Guidelines.

Irani, Z.A., Kerkhoven, E., Shojaosadati, S.A., Nielsen, J., 2016. Genome-scale metabolic model of *Pichia pastoris* with native and humanized glycosylation of recombinant proteins. Biotechnol. Bioeng. 113, 961–969. http://dx.doi.org/10.1002/bit.25863.

Iwashima, A., Nishino, H., Nose, Y., 1973. Carrier-mediated transport of thiamine in baker's yeast. Biochim. Biophys. Acta 330, 222–234. http://dx.doi.org/10.1016/0005-2736(73)90227-7.

Jacobs, P.P., Geysens, S., Vervecken, W., Contreras, R., Callewaert, N., 2009. Engineering complex-type *N*-glycosylation in *Pichia pastoris* using GlycoSwitch technology. Nat. Protoc. 4, 58–70. http://dx.doi.org/10.1038/nprot.2008.213.

Jakočiūnas, T., Bonde, I., Herrgård, M., Harrison, S.J., Kristensen, M., Pedersen, L.E., Jensen, M.K., Keasling, J.D., 2015. Multiplex metabolic pathway engineering using CRISPR/Cas9 in *Saccharomyces cerevisiae*. Metab. Eng. 28, 213–222. http://dx.doi.org/10.1016/j.ymben.2015.01.008.

Jensen, E.D., Ferreira, R., Jakočiūnas, T., Arsovska, D., Zhang, J., Ding, L., Smith, J.D., David, F., Nielsen, J., Jensen, M.K., Keasling, J.D., 2017. Transcriptional reprogramming in yeast using dCas9 and combinatorial gRNA strategies. Microb. Cell Factories 1–16. http://dx.doi.org/10.1186/s12934-017-0664-2.

Jeong, E., Shim, W.Y., Kim, J.H., 2014. Metabolic engineering of *Pichia pastoris* for production of hyaluronic acid with high molecular weight. J. Biotechnol. 185, 28–36. http://dx.doi.org/10.1016/j.jbiotec.2014.05.018.

Jeong, Y., Kim, J.-N., Kim, M.W., Bucca, G., Cho, S., Yoon, Y.J., Kim, B.-G., Roe, J.-H., Kim, S.C., Smith, C.P., Cho, B.-K., 2016. The dynamic transcriptional and translational landscape of the model antibiotic producer *Streptomyces coelicolor* A3(2). Nat. Commun. 7, 11605. http://dx.doi.org/10.1038/ncomms11605.

Jiang, M., Guo, B., Wan, X., Gong, Y., Zhang, Y., Hu, C., 2014. Isolation and characterization of the diatom Phaeodactylum delta5-elongase gene for transgenic LC-PUFA production in *Pichia pastoris*. Mar. Drugs 12, 1317–1334. http://dx.doi.org/10.3390/md12031317.

Jiang, B., Argyros, R., Bukowski, J., Nelson, S., Sharkey, N., Kim, S., Copeland, V., Davidson, R.C., Chen, R., Zhuang, J., Sethuraman, N., Stadheim, T.A., 2015. Inactivation of a GAL4-like transcription factor improves cell fitness and product yield in glycoengineered *Pichia pastoris* strains. Appl. Environ. Microbiol. 81, 260–271. http://dx.doi.org/10.1128/AEM.02619-14.

Johnson, M.A., Waterham, H.R., Ksheminska, G.P., Fayura, L.R., Cereghino, J.L., Stasyk, O.V., Veenhuis, M., Kulachkovsky, A.R., Sibirny, A.A., Cregg, J.M., 1999. Positive selection of novel peroxisome biogenesis-defective mutants of the yeast Pichia Pastoris. Genetics 151, 1379–1391.

Johnston, M., Davis, R.W., 1984. Sequences that regulate the divergent GAL1-GAL10 promoter in *Saccharomyces cerevisiae*. Mol. Cell. Biol. 4, 1440–1448. http://dx.doi.org/10.1128/MCB.4.8.1440.

Jordà, J., Suarez, C., Carnicer, M., ten Pierick, A., Heijnen, J.J., van Gulik, W., Ferrer, P., Albiol, J., Wahl, A., 2013. Glucose-methanol co-utilization in *Pichia pastoris* studied by metabolomics and instationary ^{13}C flux analysis. BMC Syst. Biol. 7, 17. http://dx.doi.org/10.1186/1752-0509-7-17.

Jordà, J., Rojas, H.C., Carnicer, M., Wahl, A., Ferrer, P., Albiol, J., 2014. Quantitative metabolomics and instationary 13 C-metabolic flux analysis reveals impact of recombinant protein production on trehalose and energy metabolism in *Pichia pastoris*. Meta 4, 281–299. http://dx.doi.org/10.3390/metabo4020281.

Kant, H.R., Balamurali, M., Meenakshisundaram, S., 2014. Enhancing precursors availability in *Pichia pastoris* for the overproduction of S-adenosyl-L-methionine employing molecular strategies with process tuning. J. Biotechnol. 188, 112–121. http://dx.doi.org/10.1016/j.jbiotec.2014.08.017.

Keasling, J.D., 2010. Manufacturing molecules through metabolic engineering. Science 330, 1355–1358. http://dx.doi.org/10.1126/science.1193990.

Karaoglan, M., Karaoglan, F.E., Inan, M., 2016. Comparison of ADH3 promoter with commonly used promoters for recombinant protein production in *Pichia pastoris*. Protein Expr. Purif. 121, 112–117. http://dx.doi.org/10.1016/j.pep.2016.01.017.

Kegel, A., Martinez, P., Carter, S.D., Åström, S.U., 2006. Genome wide distribution of illegitimate recombination events in *Kluyveromyces lactis*. Nucleic Acids Res. 34, 1633–1645. http://dx.doi.org/10.1093/nar/gkl064.

Kelwick, R., MacDonald, J.T., Webb, A.J., Freemont, P., 2014. Developments in the tools and methodologies of synthetic biology. Front. Bioeng. Biotechnol. 2, 60. http://dx.doi.org/10.3389/fbioe.2014.00060.

Khalil, A.S., Lu, T.K., Bashor, C.J., Ramirez, C.L., Pyenson, N.C., Joung, J.K., Collins, J.J., 2012. A synthetic biology framework for programming eukaryotic transcription functions. Cell 150, 647–658. http://dx.doi.org/10.1016/j.cell.2012.05.045.

Kim, S.H., Roh, K.H., Kim, K.-S., Kim, H.U., Lee, K.-R., Kang, H.-C., Kim, J.-B., 2014. Coexpression of multiple genes reconstitutes two pathways of very long-chain polyunsaturated fatty acid biosynthesis in *Pichia pastoris*. Biotechnol. Lett. 36, 1843–1851. http://dx.doi.org/10.1007/s10529-014-1550-1.

Kim, Y.J., Zhang, D., Yang, D.C., 2015. Biosynthesis and biotechnological production of ginsenosides. Biotechnol. Adv. 33, 717–735. http://dx.doi.org/10.1016/j.biotechadv.2015.03.001.

King, Z.A., Lloyd, C.J., Feist, A.M., Palsson, B.O., 2015. Next-generation genome-scale models for metabolic engineering. Curr. Opin. Biotechnol. 35, 23–29. http://dx.doi.org/10.1016/j.copbio.2014.12.016.

Kohlhaw, G.B., 2003. Leucine biosynthesis in fungi: entering metabolism through the back door. Microbiol. Mol. Biol. Rev. 67, 1–15. http://dx.doi.org/10.1128/MMBR.67.1.1.

Komor, A.C., Badran, A.H., Liu, D.R., 2016. CRISPR-based technologies for the manipulation of eukaryotic genomes. Cell 168, 1–17. http://dx.doi.org/10.1016/j.cell.2016.10.044.

Krainer, F.W., Dietzsch, C., Hajek, T., Herwig, C., Spadiut, O., Glieder, A., 2012. Recombinant protein expression in *Pichia pastoris* strains with an engineered methanol utilization pathway. Microb. Cell Factories 11, 22. http://dx.doi.org/10.1186/1475-2859-11-22.

Krainer, F.W., Gmeiner, C., Neutsch, L., Windwarder, M., Pletzenauer, R., Herwig, C., Altmann, F., Glieder, A., Spadiut, O., 2013. Knockout of an endogenous mannosyltransferase increases the homogeneity of glycoproteins produced in *Pichia pastoris*. Sci Rep 3, 3279. http://dx.doi.org/10.1038/srep03279.

Krainer, F.W., Capone, S., Jäger, M., Vogl, T., Gerstmann, M., Glieder, A., Herwig, C., Spadiut, O., 2015. Optimizing cofactor availability for the production of recombinant heme peroxidase in *Pichia pastoris*. Microb. Cell Factories 14, 4. http://dx.doi.org/10.1186/s12934-014-0187-z.

Krainer, F.W., Gerstmann, M.A., Darnhofer, B., Birner-Gruenberger, R., Glieder, A., 2016. Biotechnological advances towards an enhanced peroxidase production in *Pichia pastoris*. J. Biotechnol. 233, 181–189. http://dx.doi.org/10.1016/j.jbiotec.2016.07.012.

Küberl, A., Schneider, J., Thallinger, G.G., Anderl, I., Wibberg, D., Hajek, T., Jaenicke, S., Brinkrolf, K., Goesmann, A., Szczepanowski, R., Pühler, A., Schwab, H., Glieder, A., Pichler, H., 2011. High-quality genome sequence of *Pichia pastoris* CBS7435. J. Biotechnol. 154, 312–320. http://dx.doi.org/10.1016/j.jbiotec.2011.04.014.

Kurtzman, C.P., 2009. Biotechnological strains of *Komagataella* (*Pichia*) *pastoris* are *Komagataella phaffii* as determined from multigene sequence analysis. J. Ind. Microbiol. Biotechnol. 36, 1435–1438. http://dx.doi.org/10.1007/s10295-009-0638-4.

Lamacka, M., Sajbidor, J., 1997. Ergosterol determination in *Saccharomyces cerevisiae*. Comparison of different methods. Biotechnol. Tech. 11, 723–725. http://dx.doi.org/10.1023/A:1018484015806.

Landes, N., Gasser, B., Vorauer-Uhl, K., Lhota, G., Mattanovich, D., Maurer, M., 2016. The vitamin-sensitive promoter P THI11 enables pre-defined autonomous induction of recombinant protein production in *Pichia pastoris*. Biotechnol. Bioeng. 113, 2633–2643. http://dx.doi.org/10.1002/bit.26041.

Landry, M.P., Ando, H., Chen, A.Y., Cao, J., Kottadiel, V.I., Chio, L., Yang, D., Dong, J., Lu, T.K., Strano, M.S., 2017. Single-molecule detection of protein efflux from microorganisms using fluorescent single-walled carbon nanotube sensor arrays. Nat. Nanotechnol. 102, 26–35. http://dx.doi.org/10.1038/nnano.2016.284.

Laukens, B., De Visscher, C., Callewaert, N., 2015. Engineering yeast for producing human glycoproteins: where are we now? Future Microbiol 10, 21–34. http://dx.doi.org/10.2217/fmb.14.104.

Leber, R., Zenz, R., Schröttner, K., Fuchsbichler, S., Pühringer, B., Turnowsky, F., 2001. A novel sequence element is involved in the transcriptional regulation of expression of the ERG1 (squalene epoxidase) gene in *Saccharomyces cerevisiae*. Eur. J. Biochem. 268, 914–924. http://dx.doi.org/10.1046/j.1432-1327.2001.01940.x.

Lee, S.Y., Kim, H.U., 2015. Systems strategies for developing industrial microbial strains. Nat. Biotechnol. 33, 1061–1072. http://dx.doi.org/10.1038/nbt.3365.

Lee, S.Y., Lee, D.-Y., Kim, T.Y., 2005a. Systems biotechnology for strain improvement. Trends Biotechnol. 23, 349–358. http://dx.doi.org/10.1016/j.tibtech.2005.05.003.

Lee, C.C., Williams, T.G., Wong, D.W.S., Robertson, G.H., 2005b. An episomal expression vector for screening mutant gene libraries in *Pichia pastoris*. Plasmid 54, 80–85. http://dx.doi.org/10.1016/j.plasmid.2004.12.001.

Lee, J.W., Na, D., Park, J.M., Lee, J., Choi, S., Lee, S.Y., 2012. Systems metabolic engineering of microorganisms for natural and non-natural chemicals. Nat. Chem. Biol. 8, 536–546. http://dx.doi.org/10.1038/nchembio.970.

Lee, M.E., DeLoache, W.C., Cervantes, B., Dueber, J.E., 2015. A highly characterized yeast toolkit for modular, multipart assembly. ACS Synth. Biol. 4, 975–986. http://dx.doi.org/10.1021/sb500366v.

Levin, D.E., 2011. Regulation of cell wall biogenesis in *Saccharomyces cerevisiae*: the cell wall integrity signaling pathway. Genetics 189, 1145–1175. http://dx.doi.org/10.1534/genetics.111.128264.

Li, H., Sethuraman, N., Stadheim, T.A., Zha, D., Prinz, B., Ballew, N., Bobrowicz, P., Choi, B.-K., Cook, W.J., Cukan, M., Houston-Cummings, N.R., Davidson, R., Gong, B., Hamilton, S.R., Hoopes, J.P., Jiang, Y., Kim, N., Mansfield, R., Nett, J.H., Rios, S., Strawbridge, R., Wildt, S., Gerngross, T.U., 2006. Optimization of humanized IgGs in glycoengineered *Pichia pastoris*. Nat. Biotechnol. 24, 210–215. http://dx.doi.org/10.1038/nbt1178.

Liachko, I., Dunham, M.J., 2014. An autonomously replicating sequence for use in a wide range of budding yeasts. FEMS Yeast Res. 14, 364–367. http://dx.doi.org/10.1111/1567-1364.12123.

Liachko, I., Youngblood, R.A., Tsui, K., Bubb, K.L., Queitsch, C., Raghuraman, M.K., Nislow, C., Brewer, B.J., Dunham, M.J., 2014. GC-rich DNA elements enable replication origin activity in the methylotrophic yeast *Pichia pastoris*. PLoS Genet. 10, e1004169. http://dx.doi.org/10.1371/journal.pgen.1004169.

Liang, S., Wang, B., Pan, L., Ye, Y., He, M., Han, S., Zheng, S., Wang, X., Lin, Y., 2012. Comprehensive structural annotation of *Pichia pastoris* transcriptome and the response to various carbon sources using deep paired-end RNA sequencing. BMC

Genomics 13, 738. http://dx.doi.org/10.1186/1471-2164-13-738.
Lin Cereghino, G.P., Lin Cereghino, J., Sunga, A.J., Johnson, M.A., Lim, M., Gleeson, M.A., Cregg, J.M., 2001. New selectable marker/auxotrophic host strain combinations for molecular genetic manipulation of *Pichia pastoris*. Gene 263, 159–169. http://dx.doi.org/10.1016/S0378-1119(00)00576-X.
Lin-Cereghino, G.P., Godfrey, L., de la Cruz, B.J., Johnson, S., Khuongsathiene, S., Tolstorukov, I., Yan, M., Lin-Cereghino, J., Veenhuis, M., Subramani, S., Cregg, J.M., 2006. Mxr1p, a key regulator of the methanol utilization pathway and peroxisomal genes in *Pichia pastoris*. Mol. Cell. Biol. 26, 883–897. http://dx.doi.org/10.1128/MCB.26.3.883.
Liu, H., Tan, X., Veenhius, M., McCollum, D., Cregg, J.M., 1992. An efficient screen for peroxisome-deficient mutants of *Pichia pastoris*. J. Bacteriol. 174, 4943–4951.
Liu, L., Agren, R., Bordel, S., Nielsen, J., 2010. Use of genome-scale metabolic models for understanding microbial physiology. FEBS Lett. 584, 2556–2564. http://dx.doi.org/10.1016/j.febslet.2010.04.052.
Liu, L., Redden, H., Alper, H.S., 2013. Frontiers of yeast metabolic engineering: diversifying beyond ethanol and *Saccharomyces*. Curr. Opin. Biotechnol. 24, 1023–1030. http://dx.doi.org/10.1016/j.copbio.2013.03.005.
Liu, J.K., O'Brien, E.J., Lerman, J.A., Zengler, K., Palsson, B.O., Feist, A.M., 2014a. Reconstruction and modeling protein translocation and compartmentalization in *Escherichia coli* at the genome-scale. BMC Syst. Biol. 8, 110. http://dx.doi.org/10.1186/s12918-014-0110-6.
Liu, Y., Zhu, X., Li, W., Wen, H., Gao, Y., Liu, Y., Liu, C., 2014b. Enhancing production of ergosterol in *Pichia pastoris* GS115 by over-expression of 3-hydroxy-3-methylglutaryl CoA reductase from *Glycyrrhiza uralensis*. Acta Pharm. Sin. B 4, 161–166. http://dx.doi.org/10.1016/j.apsb.2014.02.007.
Liu, X. Bin, Liu, M., Tao, X.Y., Zhang, Z.X., Wang, F.Q., Wei, D.Z., 2015. Metabolic engineering of *Pichia pastoris* for the production of dammarenediol-II. J. Biotechnol. 216, 47–55. http://dx.doi.org/10.1016/j.jbiotec.2015.10.005.
Liu, Y., Gong, X., Wang, C., Du, G., Chen, J., Kang, Z., 2016a. Production of glucaric acid from myo-inositol in engineered *Pichia pastoris*. Enzym. Microb. Technol. 91, 8–16. http://dx.doi.org/10.1016/j.enzmictec.2016.05.009.
Liu, B., Zhang, Y., Zhang, X., Yan, C., Zhang, Y., Xu, X., Zhang, W., 2016b. Discovery of a rhamnose utilization pathway and rhamnose-inducible promoters in *Pichia pastoris*. Sci Rep 6, 27352. http://dx.doi.org/10.1038/srep27352.
Looser, V., Bruhlmann, B., Bumbak, F., Stenger, C., Costa, M., Camattari, A., Fotiadis, D., Kovar, K., 2015. Cultivation strategies to enhance productivity of *Pichia pastoris*: a review. Biotechnol. Adv. 33, 1177–1193. http://dx.doi.org/10.1016/j.biotechadv.2015.05.008.
Looser, V., Lüthy, D., Straumann, M., Hecht, K., Melzoch, K., Kovar, K., 2017. Effects of glycerol supply and specific growth rate on methanol-free production of CALB by *P. pastoris*: functional characterisation of a novel promoter. Appl. Microbiol. Biotechnol. 101, 3163–3176. http://dx.doi.org/10.1007/s00253-017-8123-x.
Love, K.R., Panagiotou, V., Jiang, B., Stadheim, T.A., Love, J.C., 2010. Integrated single-cell analysis shows *Pichia pastoris* secretes protein stochastically. Biotechnol. Bioeng. 106, 319–325. http://dx.doi.org/10.1002/bit.22688.
Love, K.R., Politano, T.J., Panagiotou, V., Jiang, B., Stadheim, T.A., Love, J.C., 2012. Systematic single-cell analysis of *Pichia pastoris* reveals secretory capacity limits productivity. PLoS One 7, 1–11. http://dx.doi.org/10.1371/journal.pone.0037915.
Love, K.R., Shah, K.A., Whittaker, C.A., Wu, J., Bartlett, M.C., Ma, D., Leeson, R.L., Priest, M., Borowsky, J., Young, S.K., Love, J.C., 2016. Comparative genomics and transcriptomics of *Pichia pastoris*. BMC Genomics 17, 550. http://dx.doi.org/10.1186/s12864-016-2876-y.
Macauley-Patrick, S., Fazenda, M.L., McNeil, B., Harvey, L.M., 2005. Heterologous protein production using the *Pichia pastoris* expression system. Yeast 22http://dx.doi.org/10.1002/yea.1208. (249–170).
Mairinger, T., Steiger, M., Nocon, J., Mattanovich, D., Koellensperger, G., Hann, S., 2015. Gas chromatography-quadrupole time-of-flight mass spectrometry-based determination of isotopologue and tandem mass Isotopomer fractions of primary metabolites for ^{13}C-metabolic flux analysis. Anal. Chem. 87, 11792–11802. http://dx.doi.org/10.1021/acs.analchem.5b03173.
Mans, R., van Rossum, H.M., Wijsman, M., Backx, A., Kuijpers, N.G.A., van den Broek, M., Daran-Lapujade, P., Pronk, J.T., van Maris, A.J.A., Daran, J.-M.G., 2015. CRISPR/Cas9: a molecular Swiss army knife for simultaneous introduction of multiple genetic modifications in *Saccharomyces cerevisiae*. FEMS Yeast Res. 1–15. http://dx.doi.org/10.1093/femsyr/fov004.
Marsalek, L., Gruber, C., Altmann, F., Aleschko, M., Mattanovich, D., Gasser, B., Puxbaum, V., 2017. Disruption of genes involved in CORVET complex leads to enhanced secretion of heterologous carboxylesterase only in protease deficient *Pichia pastoris*. Biotechnol. J. 1600584. http://dx.doi.org/10.1002/biot.201600584.
Marx, H., Mattanovich, D., Sauer, M., 2008. Overexpression of the riboflavin biosynthetic pathway in *Pichia pastoris*. Microb. Cell Factories 7, 23. http://dx.doi.org/10.1186/1475-2859-7-23.
Massahi, A., Çalik, P., 2016. Endogenous signal peptides in recombinant protein production by *Pichia pastoris*: from in-silico analysis to fermentation. J. Theor. Biol. 408, 22–33. http://dx.doi.org/10.1016/j.jtbi.2016.07.039.
Mattanovich, D., Graf, A., Stadlmann, J., Dragosits, M., Redl, A., Maurer, M., Kleinheinz, M., Sauer, M., Altmann, F., Gasser, B., 2009. Genome, secretome and glucose transport highlight unique features of the protein production host *Pichia pastoris*. Microb. Cell Factories 8, 29. http://dx.doi.org/10.1186/1475-2859-8-29.
McIsaac, R.S., Oakes, B.L., Wang, X., Dummit, K.A., Botstein, D., Noyes, M.B., 2013. Synthetic gene expression perturbation systems with rapid, tunable, single-gene specificity in yeast. Nucleic Acids Res. 41, 1–10. http://dx.doi.org/10.1093/nar/gks1313.
Meehl, M.A., Stadheim, T.A., 2014. Biopharmaceutical discovery and production in yeast. Curr. Opin. Biotechnol. 30, 120–127. http://dx.doi.org/10.1016/j.copbio.2014.06.007.
Meesapyodsuk, D., Chen, Y., Ng, S.H., Chen, J., Qiu, X., 2015. Metabolic engineering of *Pichia pastoris* to produce ricinoleic acid, a hydroxy fatty acid of industrial importance. J. Lipid Res. 56, 2102–2109. http://dx.doi.org/10.1128/AEM.02827-12.
Meyer, V., Arentshorst, M., El-Ghezal, A., Drews, A.-C., Kooistra, R., van den Hondel, C.A.M.J.J., Ram, A.F.J., 2007. Highly efficient gene targeting in the *Aspergillus niger kusA* mutant. J. Biotechnol. 128, 770–775. http://dx.doi.org/10.1016/j.jbiotec.2006.12.021.
Moreira de Almeida, J.R., Pepe de Moraes, L.M., Gonçalves Torres, F.A., 2005. Molecular characterization of the 3-phosphoglycerate kinase gene (PGK1) from the methylotrophic yeast *Pichia pastoris*. Yeast 22, 725–737. http://dx.doi.org/10.1002/yea.1243.
Näätsaari, L., Mistlberger, B., Ruth, C., Hajek, T., Hartner, F.S., Glieder, A., 2012. Deletion of the *Pichia pastoris* KU70 homologue facilitates platform strain generation for gene expression and synthetic biology. PLoS One 7, e39720. http://dx.doi.org/10.1371/journal.pone.0039720.
Nag, D.K., White, M.A., Petes, T.D., 1989. Palindromic sequences in heteroduplex DNA inhibit mismatch repair in yeast. Nature 340, 318–320. http://dx.doi.org/10.1038/340318a0.
Nett, J.H., Hodel, N., Rausch, S., Wildt, S., 2005. Cloning and disruption of the *Pichia pastoris* ARG1, ARG2, ARG3, HIS1, HIS2, HIS5, HIS6 genes and their use as auxotrophic markers. Yeast 22, 295–304. http://dx.doi.org/10.1002/yea.1202.
Nguyen, P.H., Hasek, J., Kohlwein, S.D., Romero, C., Choi, J.H., Vancura, A., 2005. Interaction of Pik1p and Sjl proteins in membrane trafficking. FEMS Yeast Res. 5, 363–371. http://dx.doi.org/10.1016/j.femsyr.2004.09.007.
Nielsen, D.R., Leonard, E., Yoon, S.H., Tseng, H.C., Yuan, C., Prather, K.L.J., 2009. Engineering alternative butanol production platforms in heterologous bacteria. Metab. Eng. 11, 262–273. http://dx.doi.org/10.1016/j.ymben.2009.05.003.
Nishida, K., Arazoe, T., Yachie, N., Banno, S., Kakimoto, M., Tabata, M., Mochizuki, M., Miyabe, A., Araki, M., Hara, K.Y., Shimatani, Z., Kondo, A., 2016. Targeted nucleotide editing using hybrid prokaryotic and vertebrate adaptive immune systems. Science 102, 553–563. http://dx.doi.org/10.1126/science.aaf8729.
Nocon, J., Steiger, M.G., Pfeffer, M., Sohn, S.B., Kim, T.Y., Maurer, M., Rußmayer, H., Pflügl, S., Ask, M., Haberhauer-Troyer, C., Ortmayr, K., Hann, S., Koellensperger, G., Gasser, B., Lee, S.Y., Mattanovich, D., 2014. Model based engineering of *Pichia pastoris* central metabolism enhances recombinant protein production. Metab. Eng. 24, 129–138. http://dx.doi.org/10.1016/j.ymben.2014.05.011.
Nocon, J., Steiger, M., Mairinger, T., Hohlweg, J., Rußmayer, H., Hann, S., Gasser, B., Mattanovich, D., 2016. Increasing pentose phosphate pathway flux enhances recombinant protein production in *Pichia pastoris*. Appl. Microbiol. Biotechnol. 100, 5955–5963. http://dx.doi.org/10.1007/s00253-016-7363-5.
Öberg, F., Sjöhamn, J., Conner, M.T., Bill, R.M., Hedfalk, K., 2011. Improving recombinant eukaryotic membrane protein yields in *Pichia pastoris*: the importance of codon optimization and clone selection. Mol. Membr. Biol. 28, 398–411. http://dx.doi.org/10.3109/09687688.2011.602219.
Obst, U., Lu, T.K., Sieber, V., 2017. A modular toolkit for generating *Pichia pastoris* secretion libraries. ACS Synth. Biol. http://dx.doi.org/10.1021/acssynbio.6b00337.
Ohsawa, S., Yurimoto, H., Sakai, Y., 2017. Novel function of Wsc proteins as a methanol-sensing machinery in the yeast *Pichia pastoris*. Mol. Microbiol. 1–47. http://dx.doi.org/10.1111/mmi.13631.
Österlund, T., Nookaew, I., Bordel, S., Nielsen, J., 2013. Mapping condition-dependent regulation of metabolism in yeast through genome-scale modeling. BMC Syst. Biol. 7, 36. http://dx.doi.org/10.1186/1752-0509-7-36.
Ostry, V., Malir, F., Ruprich, J., 2013. Producers and important dietary sources of ochratoxin A and citrinin. Toxins (Basel). 5, 1574–1586. http://dx.doi.org/10.3390/toxins5091574.
Otero, J.M., Nielsen, J., 2010. Industrial systems biology. Biotechnol. Bioeng. 105, 439–460. http://dx.doi.org/10.1002/bit.22592.
Palsson, B.O., 2006. Systems Biology - Properties of Reconstructed Networks. Cambridge University Press, Cambridge, UK.
Pan, R., Zhang, J., Shen, W.L., Tao, Z.Q., Li, S.P., Yan, X., 2011. Sequential deletion of *Pichia pastoris* genes by a self-excisable cassette. FEMS Yeast Res. 11, 292–298. http://dx.doi.org/10.1111/j.1567-1364.2011.00716.x.
Partov, S., Siewers, V., Bjorn, S., Nielsen, J., Maury, J., 2010. Characterization of different promoters for designing a new expression vector in *Saccharomyces cerevisiae*. Yeast 27, 955–964. http://dx.doi.org/10.1002/yea.1806.
Payne, M.S., Petrillo, K.L., Gavagan, J.E., Wagner, L.W., DiCosimo, R., Anton, D.L., 1995. High-level production of spinach glycolate oxidase in the methylotrophic yeast *Pichia pastoris*: engineering a biocatalyst. Gene 167, 215–219. http://dx.doi.org/10.1016/0378-1119(95)00661-3.
Payne, M.S., Petrillo, K.L., Gavagan, J.E., Dicosimo, R., Wagner, L.W., Anton, D.L., 1997. Engineering *Pichia pastoris* for biocatalysis: co-production of two active enzymes. Gene 194, 179–182. http://dx.doi.org/10.1016/S0378-1119(97)00120-0.
Peplowska, K., Markgraf, D.F., Ostrowicz, C.W., Bange, G., Ungermann, C., 2007. The CORVET tethering complex interacts with the yeast Rab5 homolog Vps21 and is involved in endo-lysosomal biogenesis. Dev. Cell 12, 739–750. http://dx.doi.org/10.1016/j.devcel.2007.03.006.
Perez-Pinera, P., Han, N., Cleto, S., Cao, J., Purcell, O., Shah, K.A., Lee, K., Ram, R., Lu, T.K., 2016. Synthetic biology and microbioreactor platforms for programmable production of biologics at the point-of-care. Nat. Commun. 7, 12211. http://dx.doi.org/10.1038/ncomms12211.
Pfeffer, M., Maurer, M., Köllensperger, G., Hann, S., Graf, A.B., Mattanovich, D., 2011. Modeling and measuring intracellular fluxes of secreted recombinant protein in *Pichia pastoris* with a novel 34S labeling procedure. Microb. Cell Factories 10, 47. http://dx.doi.org/10.1186/1475-2859-10-47.
Pfeffer, M., Maurer, M., Stadlmann, J., Grass, J., Delic, M., Altmann, F., Mattanovich, D.,

2012. Intracellular interactome of secreted antibody Fab fragment in *Pichia pastoris* reveals its routes of secretion and degradation. Appl. Microbiol. Biotechnol. 93, 2503–2512. http://dx.doi.org/10.1007/s00253-012-3933-3.
Poirier, Y., Erard, N., Petétot, J.M.C., 2002. Synthesis of polyhydroxyalkanoate in the peroxisome of *Pichia pastoris*. FEMS Microbiol. Lett. 207, 97–102. http://dx.doi.org/10.1016/S0378-1097(01)00567-5.
Poirier, Y., Antonenkov, V.D., Glumoff, T., Hiltunen, J.K., 2006. Peroxisomal β-oxidation-A metabolic pathway with multiple functions. Biochim. Biophys. Acta, Mol. Cell Res. 1763, 1413–1426. http://dx.doi.org/10.1016/j.bbamcr.2006.08.034.
Polupanov, A.S., Nazarko, V.Y., Sibirny, A.A., 2011. CCZ1, MON1 and YPT7 genes are involved in pexophagy, the Cvt pathway and non-specific macroautophagy in the methylotrophic yeast *Pichia pastoris*. Cell Biol. Int. 35, 311–319. http://dx.doi.org/10.1042/CBI20100547.
Polupanov, A.S., Nazarko, V.Y., Sibirny, A.A., 2012. Gss1 protein of the methylotrophic yeast *Pichia pastoris* is involved in glucose sensing, pexophagy and catabolite repression. Int. J. Biochem. Cell Biol. 44, 1906–1918. http://dx.doi.org/10.1016/j.biocel.2012.07.017.
Portela, R.M.C., Vogl, T., Kniely, C., Fischer, J.E., Oliveira, R., Glieder, A., 2017. Synthetic core promoters as universal parts for fine-tuning expression in different yeast species. ACS Synth. Biol. 6, 471–484. http://dx.doi.org/10.1021/acssynbio.6b00178.
Praekelt, U.M., Byrne, K.L., Meacock, P.A., 1994. Regulation of THI4(MOL1), a thiamine-biosynthetic gene of *Saccharomyces cerevisiae*. Yeast 10, 481–490. http://dx.doi.org/10.1002/yea.320100407.
Price, N.D., Reed, J.L., Palsson, B.Ø., 2004. Genome-scale models of microbial cells: evaluating the consequences of constraints. Nat. Rev. Microbiol. 2, 886–897. http://dx.doi.org/10.1038/nrmicro1023.
Prielhofer, R., Maurer, M., Klein, J., Wenger, J., Kiziak, C., Gasser, B., Mattanovich, D., 2013. Induction without methanol: novel regulated promoters enable high-level expression in *Pichia pastoris*. Microb. Cell Factories 12, 5. http://dx.doi.org/10.1186/1475-2859-12-5.
Prielhofer, R., Cartwright, S.P., Graf, A.B., Valli, M., Bill, R.M., Mattanovich, D., Gasser, B., 2015. *Pichia pastoris* regulates its gene-specific response to different carbon sources at the transcriptional, rather than the translational, level. BMC Genomics 16, 167. http://dx.doi.org/10.1186/s12864-015-1393-8.
Pronk, J.T., 2002. Auxotrophic yeast strains in fundamental and applied research. Appl. Environ. Microbiol. 68, 2095–2100. http://dx.doi.org/10.1128/AEM.68.5.2095.
Purdue, P.E., Lazarow, P.B., 2001. Peroxisome biogenesis. Annu. Rev. Cell Biol. 17, 701–752. http://dx.doi.org/10.1146/annurev.cellbio.17.1.701.
Puxbaum, V., Mattanovich, D., Gasser, B., 2015. Quo vadis? The challenges of recombinant protein folding and secretion in *Pichia pastoris*. Appl. Microbiol. Biotechnol. http://dx.doi.org/10.1007/s00253-015-6470-z.
Puxbaum, V., Gasser, B., Mattanovich, D., 2016. The bud tip is the cellular hot spot of protein secretion in yeasts. Appl. Microbiol. Biotechnol. 100, 8159–8168. http://dx.doi.org/10.1007/s00253-016-7674-6.
Qin, X., Qian, J., Yao, G., Zhuang, Y., Zhang, S., Chu, J., 2011. GAP promoter library for fine-tuning of gene expression in *Pichia pastoris*. Appl. Environ. Microbiol. 77, 3600–3608. http://dx.doi.org/10.1128/AEM.02843-10.
Raines, C.A., Lloyd, J.C., Dyer, T.A., 1999. New insights into the structure and function of but neglected Calvin cycle enzyme. J. Exp. Bot. 50, 1–8. http://dx.doi.org/10.1093/jxb/50.330.1.
Ralston, L., Subramanian, S., Matsuno, M., Yu, O., 2005. Partial reconstruction of flavonoid and isoflavonoid biosynthesis in yeast using soybean type I and type II chalcone isomerases1[w]. Plant Physiol. 137, 1375–1388. http://dx.doi.org/10.1104/pp.104.054502.1.
Rasala, B.A., Mayfield, S.P., 2015. Photosynthetic biomanufacturing in green algae; production of recombinant proteins for industrial, nutritional, and medical uses. Photosynth. Res. 123, 227–239. http://dx.doi.org/10.1007/s11120-014-9994-7.
Rebnegger, C., Graf, A.B., Valli, M., Steiger, M.G., Gasser, B., Maurer, M., Mattanovich, D., 2014. In *Pichia pastoris*, growth rate regulates protein synthesis and secretion, mating and stress response. Biotechnol. J. 9, 511–525. http://dx.doi.org/10.1002/biot.201300334.
Rebnegger, C., Vos, T., Graf, A.B., Valli, M., Pronk, J.T., Daran-Lapujade, P., Mattanovich, D., 2016. Pichia Pastoris exhibits high viability and a low maintenance energy requirement at near-zero specific growth rates. Appl. Environ. Microbiol. 82, 4570–4583. http://dx.doi.org/10.1128/AEM.00638-16.
Renuse, S., Madugundu, A.K., Kumar, P., Nair, B.G., Gowda, H., Prasad, T.S.K., Pandey, A., 2014. Proteomic analysis and genome annotation of *Pichia pastoris*, a recombinant protein expression host. Proteomics 14, 2769–2779. http://dx.doi.org/10.1002/pmic.201400267.
Ro, D.-K., Paradise, E.M., Ouellet, M., Fisher, K.J., Newman, K.L., Ndungu, J.M., Ho, K.A., Eachus, R.A., Ham, T.S., Kirby, J., Chang, M.C.Y., Withers, S.T., Shiba, Y., Sarpong, R., Keasling, J.D., 2006. Production of the antimalarial drug precursor artemisinic acid in engineered yeast. Nature 440, 3–6. http://dx.doi.org/10.1038/nature04640.
Roh, C., Seo, S.H., Choi, K.Y., Cha, M., Pandey, B.P., Kim, J.H., Park, J.S., Kim, D.H., Chang, I.S., Kim, B.G., 2009. Regioselective hydroxylation of isoflavones by *Streptomyces avermitilis* MA-4680. J. Biosci. Bioeng. 108, 41–46. http://dx.doi.org/10.1016/j.jbiosc.2009.02.021.
Roquet, N., Lu, T.K., 2014. Digital and analog gene circuits for biotechnology. Biotechnol. J. 9, 597–608. http://dx.doi.org/10.1002/biot.201300258.
Roquet, N., Soleimany, A.P., Ferris, A.C., Aaronson, S., Lu, T.K., 2016. Synthetic recombinase-based state machines in living cells. Science 353. http://dx.doi.org/10.1126/science.aad8559.
Rose, M.D., Novick, P., Thomas, J.H., Botstein, D., Fink, G.R., 1987. A *Saccharomyces cerevisiae* genomic plasmid bank based on a centromere-containing shuttle vector. Gene 60, 237–243. http://dx.doi.org/10.1016/0378-1119(87)90232-0.
Rußmayer, H., Buchetics, M., Gruber, C., Valli, M., Grillitsch, K., Modarres, G., Guerrasio, R., Klavins, K., Neubauer, S., Drexler, H., Steiger, M., Troyer, C., Al Chalabi, A., Krebiehl, G., Sonntag, D., Zellnig, G., Daum, G., Graf, A.B., Altmann, F., Koellensperger, G., Hann, S., Sauer, M., Mattanovich, D., Gasser, B., 2015a. Systems-level organization of yeast methylotrophic lifestyle. BMC Biol. 13, 80. http://dx.doi.org/10.1186/s12915-015-0186-5.
Rußmayer, H., Troyer, C., Neubauer, S., Steiger, M.G., Gasser, B., Hann, S., Koellensperger, G., Sauer, M., Mattanovich, D., 2015b. Metabolomics sampling of *Pichia pastoris* revisited: rapid filtration prevents metabolite loss during quenching. FEMS Yeast Res. 15. http://dx.doi.org/10.1093/femsyr/fov049.
Ruth, C., Buchetics, M., Vidimce, V., Kotz, D., Naschberger, S., Mattanovich, D., Pichler, H., Gasser, B., 2014. *Pichia pastoris* Aft1-a novel transcription factor, enhancing recombinant protein secretion. Microb. Cell Factories 13, 120. http://dx.doi.org/10.1186/s12934-014-0120-5.
Sánchez, B.J., Nielsen, J., 2015. Genome scale models of yeast: towards standardized evaluation and consistent omic integration. Integr. Biol. 7, 846–858. http://dx.doi.org/10.1039/C5IB00083A.
Schellenberger, J., Que, R., Fleming, R.M.T., Thiele, I., Orth, J.D., Feist, A.M., Zielinski, D.C., Bordbar, A., Lewis, N.E., Rahmanian, S., Kang, J., Hyduke, D.R., Palsson, B.Ø., 2011. Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox v2.0. Nat. Protoc. 6, 1290–1307. http://dx.doi.org/10.1038/nprot.2011.308.
Schwarzhans, J.-P., Wibberg, D., Winkler, A., Luttermann, T., Kalinowski, J., Friehs, K., 2016a. Integration event induced changes in recombinant protein productivity in *Pichia pastoris* discovered by whole genome sequencing and derived vector optimization. Microb. Cell Factories 15, 84. http://dx.doi.org/10.1186/s12934-016-0486-7.
Schwarzhans, J.-P., Wibberg, D., Winkler, A., Luttermann, T., Kalinowski, J., Friehs, K., 2016b. Non-canonical integration events in *Pichia pastoris* encountered during standard transformation analysed with genome sequencing. Sci Rep 6, 38952. http://dx.doi.org/10.1038/srep38952.
Schwarzhans, J.-P., Luttermann, T., Wibberg, D., Winkler, A., Hübner, W., Huser, T., Kalinowski, J., Friehs, K., 2017. A mitochondrial autonomously replicating sequence from *Pichia pastoris* for uniform high level recombinant protein production. Front. Microbiol. 8, 780. http://dx.doi.org/10.3389/fmicb.2017.00780.
Segrè, D., Vitkup, D., Church, G.M., 2002. Analysis of optimality in natural and perturbed metabolic networks. Proc. Natl. Acad. Sci. U. S. A. 99, 15112–15117. http://dx.doi.org/10.1073/pnas.232349399.
Shabbir Hussain, M., Rodriguez, G.M., Gao, D., Spagnuolo, M., Gambill, L., Blenner, M., 2016. Recent advances in bioengineering of the oleaginous yeast *Yarrowia lipolytica*. AIMS Bioeng. 3, 493–514. http://dx.doi.org/10.3934/bioeng.2016.4.493.
Shah, K.A., Clark, J.J., Goods, B.A., Politano, T.J., Mozdzierz, N.J., Zimnisky, R.M., Leeson, R.L., Love, J.C., Love, K.R., 2015. Automated pipeline for rapid production and screening of HIV-specific monoclonal antibodies using *Pichia pastoris*. Biotechnol. Bioeng. 112, 2624–2629. http://dx.doi.org/10.1002/bit.25663.
Shao, M., Zhang, X., Rao, Z., Xu, M., Yang, T., Li, H., Xu, Z., Yang, S., 2016. Efficient testosterone production by engineered *Pichia pastoris* co-expressing human 17β-hydroxysteroid dehydrogenase type 3 and *Saccharomyces cerevisiae* glucose 6-phosphate dehydrogenase with NADPH regeneration. Green Chem. 18, 1774–1784. http://dx.doi.org/10.1039/C5GC02353J.
Shen, S., Sulter, G., Jeffries, T.W., Cregg, J.M., 1998. A strong nitrogen source-regulated promoter for controlled expression of foreign genes in the yeast *Pichia pastoris*. Gene 216, 93–102. http://dx.doi.org/10.1016/S0378-1119(98)00315-1.
Shen, W., Kong, C., Xue, Y., Liu, Y., Cai, M., Zhang, Y., Jiang, T., Zhou, X., Zhou, M., 2016a. Kinase screening in *Pichia pastoris* identified promising targets involved in cell growth and alcohol oxidase 1 promoter (PAOX1) regulation. PLoS One 11, e0167766. http://dx.doi.org/10.1371/journal.pone.0167766.
Shen, W., Xue, Y., Liu, Y., Kong, C., Wang, X., Huang, M., Cai, M., Zhou, X., Zhang, Y., Zhou, M., 2016b. A novel methanol-free *Pichia pastoris* system for recombinant protein expression. Microb. Cell Factories 15, 178. http://dx.doi.org/10.1186/s12934-016-0578-4.
Sheng, J., Stevens, J., Feng, X., 2016. Pathway compartmentalization in peroxisome of *Saccharomyces cerevisiae* to produce versatile medium chain fatty alcohols. Sci Rep 6, 26884. http://dx.doi.org/10.1038/srep26884.
Shin, S.K., Hyeon, J.E., Kim, Y.I., Kang, D.H., Kim, S.W., Park, C., Han, S.O., 2015. Enhanced hydrolysis of lignocellulosic biomass: bi-functional enzyme complexes expressed in *Pichia pastoris* improve bioethanol production from *Miscanthus sinensis*. Biotechnol. J. 10, 1912–1919. http://dx.doi.org/10.1002/biot.201500081.
Shoji, J., Kikuma, T., Kitamoto, K., 2014. Vesicle trafficking, organelle functions, and unconventional secretion in fungal physiology and pathogenicity. Curr. Opin. Microbiol. 20, 1–9. http://dx.doi.org/10.1016/j.mib.2014.03.002.
Sohn, S.B., Graf, A.B., Kim, T.Y., Gasser, B., Maurer, M., Ferrer, P., Mattanovich, D., Lee, S.Y., 2010. Genome-scale metabolic model of methylotrophic yeast *Pichia pastoris* and its use for in silico analysis of heterologous protein production. Biotechnol. J. 5, 705–715. http://dx.doi.org/10.1002/biot.201000078.
Solà, A., Jouhten, P., Maaheimo, H., Sánchez-Ferrando, F., Szyperski, T., Ferrer, P., 2007. Metabolic flux profiling of *Pichia pastoris* grown on glycerol/methanol mixtures in chemostat cultures at low and high dilution rates. Microbiology 153, 281–290. http://dx.doi.org/10.1099/mic.0.29263-0.
Soma, Y., Hanai, T., 2015. Self-induced metabolic state switching by a tunable cell density sensor for microbial isopropanol production. Metab. Eng. 30, 7–15. http://dx.doi.org/10.1016/j.ymben.2015.04.005.
Spong, A.P., Subramani, S., 1993. Cloning and characterization of Pas5 - a gene required for peroxisome biogenesis in the methylotrophic yeast *Pichia Pastoris*. J. Cell Biol. 123, 535–548. http://dx.doi.org/10.1083/jcb.123.3.535.
Srivastava, A., Somvanshi, P., Mishra, B.N., 2012. Reconstruction and visualization of carbohydrate, *N*-glycosylation pathways in *Pichia pastoris* CBS7435 using computational and system biology approaches. Syst. Synth. Biol. 7, 7–22. http://dx.doi.org/

10.1007/s11693-012-9102-2.
Stadlmayr, G., Mecklenbräuker, A., Rothmüller, M., Maurer, M., Sauer, M., Mattanovich, D., Gasser, B., 2010. Identification and characterisation of novel *Pichia pastoris* promoters for heterologous protein production. J. Biotechnol. 150, 519–529. http://dx.doi.org/10.1016/j.jbiotec.2010.09.957.
Stanley, D., Bandara, A., Fraser, S., Chambers, P.J., Stanley, G.A., 2010. The ethanol stress response and ethanol tolerance of *Saccharomyces cerevisiae*. J. Appl. Microbiol. 109, 13–24. http://dx.doi.org/10.1111/j.1365-2672.2009.04657.x.
Staunton, J., Weissman, K.J., 2001. Polyketide biosynthesis: a millennium review. Nat. Prod. Rep. 18, 380–416. http://dx.doi.org/10.1039/a909079g.
Storici, F., Durham, C.L., Gordenin, D.A., Resnick, M.A., 2003. Chromosomal site-specific double-strand breaks are efficiently targeted for repair by oligonucleotides in yeast. Proc. Natl. Acad. Sci. U. S. A. 100, 14994–14999. http://dx.doi.org/10.1073/pnas.2036296100.
Sturmberger, L., Chappell, T., Geier, M., Krainer, F., Day, K.J., Vide, U., Trstenjak, S., Schiefer, A., Richardson, T., Soriaga, L., Darnhofer, B., Birner-Gruenberger, R., Glick, B.S., Tolstorukov, I., Cregg, J., Madden, K., Glieder, A., 2016. Refined *Pichia pastoris* reference genome sequence. J. Biotechnol. 235, 121–131. http://dx.doi.org/10.1016/j.jbiotec.2016.04.023.
Suzuki, K., Imai, Y., Yamashita, I., 1983. *In vivo* ligation of linear DNA molecules to circular forms in the yeast *Saccharomyces cerevisiae*. J. Bacteriol. 155, 747–754.
Symington, L.S., 2002. Role of RAD52 epistasis group genes in homologous recombination and double-strand break repair. Microbiol. Mol. Biol. Rev. 66, 630–670. http://dx.doi.org/10.1128/MMBR.66.4.630.
Szymczak, A.L., Workman, C.J., Wang, Y., Vignali, K.M., Dilioglou, S., Vanin, E.F., Vignali, D.A.A., 2004. Correction of multi-gene deficiency in vivo using a single "self-cleaving" 2A peptide-based retroviral vector. Nat. Biotechnol. 22, 589–594. http://dx.doi.org/10.1038/nbt957.
Takagi, S., Tsutsumi, N., Terui, Y., Kong, X.Y., 2012. Method for Methanol Independent Induction From Methanol Inducible Promoters in Pichia. (US8236528 B2).
Tao, H., Guo, D., Zhang, Y., Deng, Z., Liu, T., 2015. Metabolic engineering of microbes for branched-chain biodiesel production with low-temperature property. Biotechnol. Biofuels 8, 92. http://dx.doi.org/10.1186/s13068-015-0270-7.
Tomàs-Gamisans, M., Ferrer, P., Albiol, J., 2016. Integration and validation of the genome-scale metabolic models of *Pichia pastoris*: a comprehensive update of protein glycosylation pathways, lipid and energy metabolism. PLoS One 11, 1–24. http://dx.doi.org/10.1371/journal.pone.0148031.
Trantas, E., Panopoulos, N., Ververidis, F., 2009. Metabolic engineering of the complete pathway leading to heterologous biosynthesis of various flavonoids and stilbenoids in *Saccharomyces cerevisiae*. Metab. Eng. 11, 355–366. http://dx.doi.org/10.1016/j.ymben.2009.07.004.
Tredwell, G.D., Edwards-Jones, B., Leak, D.J., Bundy, J.G., 2011. The development of metabolomic sampling procedures for *Pichia pastoris*, and baseline metabolome data. PLoS One 6. http://dx.doi.org/10.1371/journal.pone.0016286.
Tredwell, G.D., Aw, R., Edwards-Jones, B., Leak, D.J., Bundy, J.G., 2017. Rapid screening of cellular stress responses in recombinant *Pichia pastoris* strains using metabolite profiling. J. Ind. Microbiol. Biotechnol. 44, 413–417. http://dx.doi.org/10.1007/s10295-017-1904-5.
Trinklein, N.T., Force Aldred, S., Hartman, S.J., Schroeder, D.I., Otillar, R.P., Myers, R.M., 2004. An abundance of bidirectional promoters in the human genome. Genome Res. 14, 62–66. http://dx.doi.org/10.1101/gr.1982804.
Tsakraklides, V., Brevnova, E., Stephanopoulos, G., Shaw, A.J., 2015. Improved gene targeting through cell cycle synchronization. PLoS One 10, 1–11. http://dx.doi.org/10.1371/journal.pone.0133434.
Tschopp, J.F., Brust, P.F., Cregg, J.M., Stillman, C.A., Gingeras, T.R., 1987. Expression of the lacZ gene from two methanol-regulated promoters in *Pichia pastoris*. Nucleic Acids Res. 15, 3859–3876.
Valli, M., Tatto, N.E., Peymann, A., Gruber, C., Landes, N., Ekker, H., Thallinger, G.G., Mattanovich, D., Gasser, B., Graf, A.B., 2016. Curation of the genome annotation of *Pichia pastoris* (*Komagataella phaffii*) CBS7435 from gene level to protein function. FEMS Yeast Res. 1–28. http://dx.doi.org/10.1093/femsyr/fow051.
van der Klei, I.J., Yurimoto, H., Sakai, Y., Veenhuis, M., 2006. The significance of peroxisomes in methanol metabolism in methylotrophic yeast. Biochim. Biophys. Acta, Mol. Cell Res. 1763, 1453–1462. http://dx.doi.org/10.1016/j.bbamcr.2006.07.016.
Van Dijk, E.L., Auger, H., Jaszczyszyn, Y., Thermes, C., 2014. Ten years of next-generation sequencing technology. Trends Genet. 30, 418–426. http://dx.doi.org/10.1016/j.tig.2014.07.001.
Vassileva, A., Chugh, D.A., Swaminathan, S., Khanna, N., 2001. Effect of copy number on the expression levels of hepatitis B surface antigen in the methylotrophic yeast *Pichia pastoris*. Protein Expr. Purif. 21, 71–80. http://dx.doi.org/10.1006/prep.2000.1335.
Verbeke, J., Beopoulos, A., Nicaud, J.M., 2013. Efficient homologous recombination with short length flanking fragments in Ku70 deficient *Yarrowia lipolytica* strains. Biotechnol. Lett. 35, 571–576. http://dx.doi.org/10.1007/s10529-012-1107-0.
Verna, J., Lodder, A., Lee, K., Vagts, A., Ballester, R., 1997. A family of genes required for maintenance of cell wall integrity and for the stress response in *Saccharomyces cerevisiae*. Proc. Natl. Acad. Sci. U. S. A. 94, 13804–13809. http://dx.doi.org/10.1073/pnas.94.25.13804.
Vervecken, W., Kaigorodov, V., Callewaert, N., Geysens, S., De Vusser, K., Contreras, R., 2004. In vivo synthesis of mammalian-like, hybrid-type *N*-glycans in *Pichia pastoris*. Appl. Environ. Microbiol. 70, 2639–2646. http://dx.doi.org/10.1128/AEM.70.5.2639-2646.2004.
Verwaal, R., Wang, J., Meijnen, J.P., Visser, H., Sandmann, G., Van Den Berg, J.A., Van Ooyen, A.J.J., 2007. High-level production of beta-carotene in *Saccharomyces cerevisiae* by successive transformation with carotenogenic genes from *Xanthophyllomyces dendrorhous*. Appl. Environ. Microbiol. 73, 4342–4350. http://dx.doi.org/10.1128/AEM.02759-06.
Vogl, T., Glieder, A., 2013. Regulation of *Pichia pastoris* promoters and its consequences for protein production. New Biotechnol. 30, 385–404. http://dx.doi.org/10.1016/j.nbt.2012.11.010.
Vogl, T., Ruth, C., Pitzer, J., Kickenweiz, T., Glieder, A., 2014. Synthetic core promoters for *Pichia pastoris*. ACS Synth. Biol. 3, 188–191. http://dx.doi.org/10.1021/sb400091p.
Vogl, T., Ahmad, M., Krainer, F.W., Schwab, H., Glieder, A., 2015a. Restriction site free cloning (RSFC) plasmid family for seamless, sequence independent cloning in *Pichia pastoris*. Microb. Cell Factories 14, 103. http://dx.doi.org/10.1186/s12934-015-0293-6.
Vogl, T., Kickenweiz, T., Sturmberger, L., Glieder, A., 2015b. Bidirectional Promoter. EP2862933A3.
Vogl, T., Sturmberger, L., Kickenweiz, T., Wasmayer, R., Schmid, C., Hatzl, A.-M., Gerstmann, M.A., Pitzer, J., Wagner, M., Thallinger, G.G., Geier, M., Glieder, A., 2016. A toolbox of diverse promoters related to methanol utilization – functionally verified parts for heterologous pathway expression in *Pichia pastoris*. ACS Synth. Biol. 5, 172–186. http://dx.doi.org/10.1021/acssynbio.5b00199.
Wang, S., Yao, Q., Tao, J., Qiao, Y., Zhang, Z., 2007. Co-ordinate expression of glycine betaine synthesis genes linked by the FMDV 2A region in a single open reading frame in *Pichia pastoris*. Appl. Microbiol. Biotechnol. 77, 891–899. http://dx.doi.org/10.1007/s00253-007-1222-3.
Wang, T.Y., Tsai, Y.H., Yu, I.Z., Chang, T.S., 2015. Improving 3′-hydroxygenistein production in recombinant *Pichia pastoris* using periodic hydrogen peroxide-shocking strategy. J. Microbiol. Biotechnol. 26, 498–502. http://dx.doi.org/10.4014/jmb.1509.09013.
Wang, X., Wang, Q., Wang, J., Bai, P., Shi, L., Shen, W., Zhou, M., Zhou, X., Zhang, Y., Cai, M., 2016. Mit1 transcription factor mediates methanol signaling and regulates alcohol oxidase 1 promoter in *Pichia pastoris*. J. Biol. Chem. 291, 6245–6261. http://dx.doi.org/10.1074/jbc.M115.692053.
Wang, J., Wang, X., Shi, L., Qi, F., Zhang, P., Zhang, Y., Zhou, X., Song, Z., Cai, M., 2017. Methanol-independent protein expression by AOX1 promoter with trans-acting elements engineering and glucose-glycerol-shift induction in *Pichia pastoris*. Sci Rep 7, 41850. http://dx.doi.org/10.1038/srep41850.
Waterham, H.R., Cregg, J.M., 1997. Peroxisome biogenesis. BioEssays 19, 57–66. http://dx.doi.org/10.1002/bies.950190110.
Waterham, H.R., Digan, M.E., Koutz, P.J., Lair, S.V., Cregg, J.M., 1997a. Isolation of the *Pichia pastoris* glyceraldehyde-3-phosphate dehydrogenase gene and regulation and use of its promoter. Gene 186, 37–44.
Waterham, H.R., Russell, K.A., De Vries, Y., Cregg, J.M., 1997b. Peroxisomal targeting, import, and assembly of alcohol oxidase in *Pichia pastoris*. J. Cell Biol. 139, 1419–1431. http://dx.doi.org/10.1083/jcb.139.6.1419.
Weinhandl, K., Ballach, M., Winkler, M., Ahmad, M., Glieder, A., Birner-Gruenberger, R., Fotheringham, I., Escalettes, F., Camattari, A., 2016. *Pichia pastoris* mutants as host strains for efficient secretion of recombinant branched chain aminotransferase (BCAT). J. Biotechnol. 235, 84–91. http://dx.doi.org/10.1016/j.jbiotec.2016.06.004.
Weninger, A., Glieder, A., Vogl, T., 2015. A toolbox of endogenous and heterologous nuclear localization sequences for the methylotrophic yeast *Pichia pastoris*. FEMS Yeast Res. 15. http://dx.doi.org/10.1093/femsyr/fov082.
Weninger, A., Hatzl, A.-M., Schmid, C., Vogl, T., Glieder, A., 2016. Combinatorial optimization of CRISPR/Cas9 expression enables precision genome engineering in the methylotrophic yeast *Pichia pastoris*. J. Biotechnol. 235, 139–149. http://dx.doi.org/10.1016/j.jbiotec.2016.03.027.
Werten, M.W.T., De Wolf, F.A., 2005. Reduced proteolysis of secreted gelatin and Yps1-mediated alpha-factor leader processing in a *Pichia pastoris* kex2 disruptant. Appl. Environ. Microbiol. 71, 2310–2317. http://dx.doi.org/10.1128/AEM.71.5.2310-2317.2005.
Withers, S.T., Keasling, J.D., 2007. Biosynthesis and engineering of isoprenoid small molecules. Appl. Microbiol. Biotechnol. 73, 980–990. http://dx.doi.org/10.1007/s00253-006-0593-1.
Wriessnegger, T., Gübitz, G., Leitner, E., Ingolic, E., Cregg, J., de la Cruz, B.J., Daum, G., 2007. Lipid composition of peroxisomes from the yeast *Pichia pastoris* grown on different carbon sources. Biochim. Biophys. Acta 1771, 455–461. http://dx.doi.org/10.1016/j.bbalip.2007.01.004.
Wriessnegger, T., Augustin, P., Engleder, M., Leitner, E., Müller, M., Kaluzna, I., Schürmann, M., Mink, D., Zellnig, G., Schwab, H., Pichler, H., 2014. Production of the sesquiterpenoid (+)-nootkatone by metabolic engineering of *Pichia pastoris*. Metab. Eng. 24, 18–29. http://dx.doi.org/10.1016/j.ymben.2014.04.001.
Wriessnegger, T., Moser, S., Emmerstorfer-Augustin, A., Leitner, E., Müller, M., Kaluzna, I., Schürmann, M., Mink, D., Pichler, H., 2016. Enhancing cytochrome P450-mediated conversions in *P. pastoris* through RAD52 over-expression and optimizing the cultivation conditions. Fungal Genet. Biol. 89, 114–125. http://dx.doi.org/10.1016/j.fgb.2016.02.004.
Wu, S., Fallon, R.D., Payne, M.S., 1999. Engineering *Pichia pastoris* for stereoselective nitrile hydrolysis by co-producing three heterologous proteins. Appl. Microbiol. Biotechnol. 52, 186–190. http://dx.doi.org/10.1007/s002530051507.
Xiao, S., Gao, Y., Wang, X., Shen, W., Wang, J., Zhou, X., Cai, M., Zhang, Y., 2016. Peroxisomes targeted and tandem repeat multimer expressions of human antimicrobial peptide LL37 in *Pichia pastoris*. Prep. Biochem. Biotechnol. 1–7. http://dx.doi.org/10.1080/10826068.2016.1201684.
Xu, Z., Wei, W., Gagneur, J., Perocchi, F., Clauder-Münster, S., Camblong, J., Guffanti, E., Stutz, F., Huber, W., Steinmetz, L.M., 2009. Bidirectional promoters generate pervasive transcription in yeast. Nature 457, 1033–1037. http://dx.doi.org/10.1038/nature07728.
Xue, Y., Kong, C., Shen, W., Bai, C., Ren, Y., Zhou, X., Zhang, Y., Cai, M., 2017. Methylotrophic yeast *Pichia pastoris* as a chassis organism for polyketide synthesis via the full citrinin biosynthetic pathway. J. Biotechnol. 242, 64–72. http://dx.doi.org/

10.1016/j.jbiotec.2016.11.031.
Yan, M., Rachubinski, D.A., Joshi, S., Rachubinski, R.A., Subramani, S., 2008. Dysferlin domain-containing proteins, Pex30p and Pex31p, localized to two compartments, control the number and size of oleate-induced peroxisomes in *Pichia pastoris*. Mol. Biol. Cell 19, 885–898. http://dx.doi.org/10.1091/mbc.E07-10-1042.
Yan, J., Zheng, X., Du, L., Li, S., 2014. Integrated lipase production and in situ biodiesel synthesis in a recombinant *Pichia pastoris* yeast: an efficient dual biocatalytic system composed of cell free enzymes and whole cell catalysts. Biotechnol. Biofuels 7, 55. http://dx.doi.org/10.1186/1754-6834-7-55.
Yurimoto, H., Oku, M., Sakai, Y., 2011. Yeast Methylotrophy: Metabolism, Gene Regulation and Peroxisome Homeostasis. Int. J. Microbiol. 2011, 1–8. http://dx.doi.org/10.1155/2011/101298.
Yu, X.-W., Wang, R., Zhang, M., Xu, Y., Xiao, R., 2012. Enhanced thermostability of a Rhizopus chinensis lipase by in vivo recombination in *Pichia pastoris*. Microb. Cell Factories 11, 102–112. http://dx.doi.org/10.1186/1475-2859-11-102.
Zecchin, A., Stapor, P.C., Goveia, J., Carmeliet, P., 2015. Metabolic pathway compartmentalization: an underappreciated opportunity? Curr. Opin. Biotechnol. 34, 73–81. http://dx.doi.org/10.1016/j.copbio.2014.11.022.
Zhang, P., Zhang, W., Zhou, X., Bai, P., Cregg, J.M., Zhang, Y., 2010. Catabolite repression of Aox in *Pichia pastoris* is dependent on hexose transporter PpHxt1 and pexophagy. Appl. Environ. Microbiol. 76, 6108–6118. http://dx.doi.org/10.1128/AEM.00607-10.
Zhang, R., Patena, W., Armbruster, U., Gang, S.S., Blum, S.R., Jonikas, M.C., 2014. High-throughput genotyping of green algal mutants reveals random distribution of mutagenic insertion sites and endonucleolytic cleavage of transforming DNA. Plant Cell 26, 1398–1409. http://dx.doi.org/10.1105/tpc.114.124099.
Zhang, M., Liu, Z., Yu, Q., Mao, J., Zhang, B., Xing, L., Li, M., 2015. Deletion of genes encoding fatty acid desaturases leads to alterations in stress sensitivity in *Pichia pastoris*. FEMS Yeast Res. 15. http://dx.doi.org/10.1093/femsyr/fov020.
Zhang, Y., Wang, C., Wang, L., Yang, R., Hou, P., 2017. Direct bioethanol production from wheat straw using xylose/glucose co-fermentation by co-culture of two recombinant yeasts. J. Ind. Microbiol. Biotechnol. 44, 453–464. http://dx.doi.org/10.1007/s10295-016-1893-9.
Zhao, C., Gao, X., Liu, X., Wang, Y., Yang, S., Wang, F., Ren, Y., 2016. Enhancing biosynthesis of a ginsenoside precursor by self-assembly of two key enzymes in *Pichia pastoris*. J. Agric. Food Chem. 64, 3380–3385. http://dx.doi.org/10.1021/acs.jafc.6b00650.
Zhou, Y.J., Buijs, N.A., Zhu, Z., Gómez, D.O., Boonsombuti, A., Siewers, V., Nielsen, J., 2016. Harnessing yeast peroxisomes for biosynthesis of fatty-acid-derived biofuels and chemicals with relieved side-pathway competition. J. Am. Chem. Soc. 138, 15368–15377. http://dx.doi.org/10.1021/jacs.6b07394.
Zhu, T., Guo, M., Sun, C., Qian, J., Zhuang, Y., Chu, J., Zhang, S., 2009. A systematical investigation on the genetic stability of multi-copy *Pichia pastoris* strains. Biotechnol. Lett. 31, 679–684. http://dx.doi.org/10.1007/s10529-009-9917-4.
Zhuang, S., Fu, J., Powell, C., Huang, J., Xia, Y., Yan, R., 2015. Production of medium-chain volatile flavour esters in *Pichia pastoris* whole-cell biocatalysts with extracellular expression of *Saccharomyces cerevisiae* acyl-CoA:ethanol O-acyltransferase Eht1 or Eeb1 background. Spring 4. http://dx.doi.org/10.1186/s40064-015-1195-0.
Zirpel, B., Stehle, F., Kayser, O., 2015. Production of Delta9-tetrahydrocannabinolic acid from cannabigerolic acid by whole cells of *Pichia* (*Komagataella*) *pastoris* expressing Delta9-tetrahydrocannabinolic acid synthase from *Cannabis sativa* L. Biotechnol. Lett. 37, 1869–1875. http://dx.doi.org/10.1007/s10529-015-1853-x.