

RESEARCH

Open Access



# Scalable protein production by *Komagataella phaffii* enabled by ARS plasmids and carbon source-based selection

Florian Weiss<sup>1</sup>, Guillermo Requena-Moreno<sup>2</sup>, Carsten Pichler<sup>1</sup>, Francisco Valero<sup>2</sup>, Anton Glieder<sup>1\*</sup> and Xavier Garcia-Ortega<sup>2</sup>

## Abstract

**Background** Most recombinant *Komagataella phaffii* (*Pichia pastoris*) strains for protein production are generated by genomic integration of expression cassettes. The clonal variability in gene copy numbers, integration loci and consequently product titers limit the aptitude for high throughput applications in drug discovery, enzyme engineering or most comparative analyses of genetic elements such as promoters or secretion signals. Circular episomal plasmids with an autonomously replicating sequence (ARS), an alternative which would alleviate some of these limitations, are inherently unstable in *K. phaffii*. Permanent selection pressure, mostly enabled by antibiotic resistance or auxotrophy markers, is crucial for plasmid maintenance and hardly scalable for production. The establishment and use of extrachromosomal ARS plasmids with key genes of the glycerol metabolism (glycerol kinase 1, *GUT1*, and triosephosphate isomerase 1, *TPI1*) as selection markers was investigated to obtain a system with high transformation rates that can be directly used for scalable production processes in lab scale bioreactors.

**Results** In micro-scale deep-well plate experiments, ARS plasmids employing the *Ashbya gossypii* *TEF1* (transcription elongation factor 1) promoter to regulate transcription of the marker gene were found to deliver high transformation efficiencies and the best performances with the reporter protein (*CalB*, lipase B of *Candida antarctica*) for both, the *GUT1*- and *TPI1*-based, marker systems. The *GUT1* marker-bearing strain surpassed the reference strain with integrated expression cassette by 46% upon re-evaluation in shake flask cultures regarding *CalB* production, while the *TPI1* system was slightly less productive compared to the control. In 5 L bioreactor methanol-free fed-batch cultivations, the episomal production system employing the *GUT1* marker led to 100% increased *CalB* activity in the culture supernatant compared to integration construct.

**Conclusions** For the first time, a scalable and methanol-independent expression system for recombinant protein production for *K. phaffii* using episomal expression vectors was demonstrated. Expression of the *GUT1* selection marker gene of the new ARS plasmids was refined by employing the *TEF1* promoter of *A. gossypii*. Additionally, the antibiotic-free marker toolbox for *K. phaffii* was expanded by the *TPI1* marker system, which proved to be similarly suited for the use in episomal plasmids as well as integrative expression constructs for the purpose of recombinant protein production.

\*Correspondence:  
Anton Glieder  
a.glieder@tugraz.at

Full list of author information is available at the end of the article



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

**Keywords** *Komagataella phaffii*, *Pichia pastoris*, Episomal ARS plasmid, Lipase CalB, Carbon source marker, *GUT1*, *TPI1*, Methanol-free bioprocesses

## Background

The methylotrophic yeast *Komagataella phaffii*, formerly known as *Pichia pastoris*, is a preferred microbial host for the production of recombinant proteins and metabolites in research and industrial applications [1]. Typical eukaryotic post-translational modifications enable the production of proteins which are otherwise difficult to synthesize or fold by more simple microbial hosts [2]. Most production strains are created by integration of the expression vector into the genome [3, 4], often using high DNA concentrations and increased selection pressure to support multi-copy integration. While this usually results in genetically stable clones at least for single copy strains, non-specific integration - of potentially multiple (often tandem) copies - might positively or adversely influence heterologous gene expression or protein secretion due to the affected genomic loci where the plasmid integrated via non-homologous end-joining (NHEJ) [5–7]. Multiple copies of similar expression cassettes integrated in close proximity are usually not genetically stable [8].

Certain applications like high-throughput library construction and screening or direct comparison of enzyme/promoter/secretion signal variants, however, would benefit from more efficient transformation and more uniform expression levels of different transformants by reduction of the widely described clonal variability [9, 10]. However, due to the low efficiency of homologous recombination in *K. phaffii*, so far transformation rates with linearized plasmids for site-directed integration were usually limited to max.  $10^3 - 10^4$  transformants per  $\mu\text{g}$  DNA - not enough for applications where high rates of transformation are desired [11]. The use of circular autonomously replicating episomal plasmids, on the other hand, has been shown to drastically increase transformation rates [12, 13]. In addition, the above-mentioned locus effects caused by integrative vectors might be alleviated [12]. In absence of a functional 2-micron plasmid for *K. phaffii*, such plasmids usually carry an autonomously replicating sequence (ARS) that serves as an origin of replication. Besides endogenous sequences like the well-studied *Pichia* ARS 1 and 2 (PARS1 and PARS2) and a mitochondrial ARS, sequences from *Candida boidinii* and *Kluyveromyces lactis* have been described and successfully implemented in *K. phaffii* so far [12–15]. But unlike the well-investigated 2-micron plasmids in *Saccharomyces cerevisiae* [16, 17], currently available ARS plasmid systems in *K. phaffii* were inherently unstable and require constant selection in order to be maintained [9, 10, 18]. After initial work with ARS-based plasmids by Cregg et al. [12], those vectors were given up and integrative

vectors were preferred due to higher stability of transformants and more reliable expression. About 20 years later the *HIS4* selection marker was replaced by the simpler and more selective Zeocin resistance marker [9] but also this was not leading to any breakthrough and no production processes employing episomal plasmid vectors were reported so far. Due to the inherent instability of all early-generation episomal plasmids, efforts were made to facilitate mitotic distribution by placing centromeric sequences on the plasmids [18, 19]. This seemingly stabilized the plasmids to approximately one copy per cell, but their large size and low transformation efficiencies compared to standard ARS plasmids still restrict their use for the desired applications in high throughput experiments. For all types of plasmids selection pressure was usually applied by the use of antibiotic or auxotrophy markers [9, 10, 14, 20]. While antibiotic resistance genes are convenient dominant markers in small-scale lab experiments, their use at large scale is not suited for most product applications and is often considered economically unfeasible [21]. Prototrophic markers, on the other hand, complement deficiencies of auxotrophic host strains and therefore provide advantages for the host cells on defined media [21]. This selection pressure, however, can be diminished by cross-feeding during the cultivation, potentially leading to marker-less subpopulations [22]. Alternatively, genes that are required for the utilization of certain carbon sources could be used as efficient and economically feasible selection markers as already demonstrated for *Schizosaccharomyces pombe* and *S. cerevisiae* over 20 years ago, reviewed by Kjeldsen et al. as well as patented by Kawasaki and Bell [23, 24]. Therefore, a similar strategy might be successful for *K. phaffii*.

For example, the *GUT1* gene, coding for a glycerol kinase essential for glycerol metabolism, has been used as a marker gene in a *gut1* deficient strain for the cultivation in defined media with glycerol as the sole carbon source by Näätsaari et al. [11], and was also implemented on ARS plasmids [13]. As a possible alternative *TPI1*, which codes for a triosephosphate isomerase that catalyzes the interconversion of dihydroxyacetone phosphate to glyceraldehyde 3-phosphate and thus essentially connects the glycerol metabolism to glycolysis, had been shown to serve as an auto-selection marker in *S. cerevisiae* [24], which means it even provides selective pressure in complex media [21]. For *K. phaffii* the *TPI1* marker had not been used so far. Both *GUT1* and *TPI1* require strains that are deficient in the respective genes, which are then complemented by the selection marker cassette present in the episomal plasmid.

The goals of this study were to evaluate and improve the *GUT1* marker system for its application in ARS plasmids and to be implemented - for the first time - in scalable bioreactor cultures. In addition, the idea was to compare this selection system to a possible new *TPII* marker system in *K. phaffii*. For direct comparison, the production of *Candida antarctica* lipase B (*CalB*), which is of industrial interest for a wide range of applications, was used as a reporter protein for both systems. Combinations of 5 distinct promoters of different regulations driving the marker gene expression were tested together with the PARS1 sequence in terms of *CalB* reporter activity in *gut1* and *tpi1* deficient strains. The latter had to be generated specifically for this study, because - to the best of our knowledge - no studies regarding a deletion of *TPII* in *K. phaffii* were published at the time of writing this publication.

## Results and discussion

### $P_{AgTEF1}$ lead to the highest reporter activities in deep-well plate scale

In order to find a reliable episomal plasmid system that is based on selection pressure provided by the carbon source in the media, the native *TPII* and a codon-optimized *GUT1* gene, both essential for glycerol metabolism, were tested in combination with 5 different promoters of various strengths on a plasmid harboring the PARS1

sequence and a *CalB* expression cassette as the reporter. All plasmids generated in this study are summarized in Table 1. After the transformation of the *tpi1* or *gut1* deficient strains (BSYBG11  $\Delta tpi1$  and BSY11G1) with the respective plasmids the transformation efficiencies were determined (Figure S1, Supplementary file 1) after 5 days of incubation at 28 °C. Most integrative plasmids ranged between 1.5 and  $3 \times 10^3$  CFU/ $\mu$ g with the exceptions of  $P_{ILV5}$  - *GUT1* where no transformants could be obtained and  $P_{GCW14}$  - *TPII* which showed a transformation efficiency of  $0.5 \times 10^3$  CFU/ $\mu$ g. On the other hand, transformation efficiencies of *TPII*-based episomal plasmids ranged from 1.0 to  $1.4 \times 10^6$  and *GUT1*-based episomal plasmids from 1.5 to  $2 \times 10^6$  CFU/ $\mu$ g DNA. Here, the  $P_{GCW14}$  - *GUT1* combination produced transformants of such small size that they were almost not visible or distinguishable from the background (Supplementary file 1, Figure S2), and therefore was incubated for one more day to obtain some colonies for inoculation. Thus, the determination of the transformation efficiency could not be done in a reliable way. These results support the information found in literature that ARS containing plasmids, given that the selection marker expression is properly adjusted, enable higher transformation rates in yeasts, or specifically *K. phaffii*, than integration vectors [12, 25].

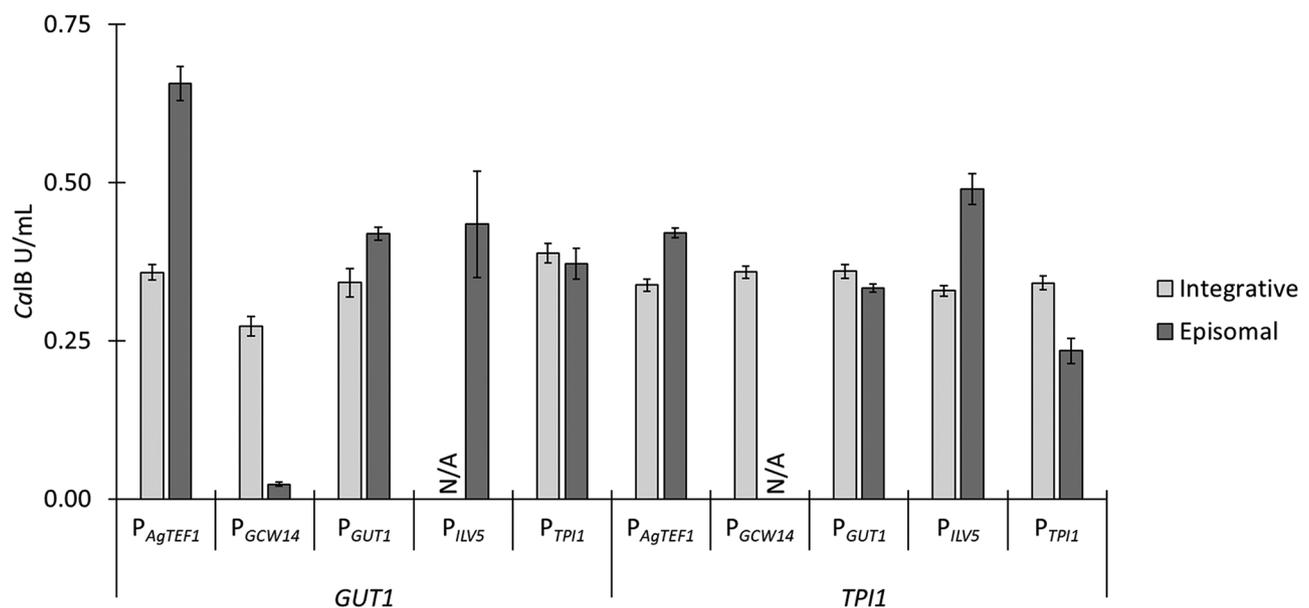
Seven transformants of each ARS plasmid (where applicable) were cultivated in 96 deep-well plates

**Table 1** All plasmids generated in this study

Plasmid Number	Description	Selection marker	Name	Promoter	PARS1
1	Cloning vectors; not used in <i>K. phaffii</i>	<i>GUT1/Ampicillin</i>	pBSY3G	$P_{GUT1}$	-
2			pBSY3G_PARS1	$P_{GUT1}$	+
3			pBSY3G_DaCalB	$P_{GUT1}$	-
4			pBSY3G_PARS1_DaCalB	$P_{GUT1}$	+
5	<i>K. phaffii</i> expression vectors	<i>GUT1</i>	pBSY3G_DaCalB_ $P_{AgTEF1}$	$P_{AgTEF1}$	-
6			pBSY3G_PARS1_DaCalB_ $P_{AgTEF1}$	$P_{AgTEF1}$	+
7			pBSY3G_DaCalB_ $P_{GCW14}$	$P_{GCW14}$	-
8			pBSY3G_PARS1_DaCalB_ $P_{GCW14}$	$P_{GCW14}$	+
9			pBSY3G_DaCalB_ $P_{GUT1}$	$P_{GUT1}$	-
10			pBSY3G_PARS1_DaCalB_ $P_{GUT1}$	$P_{GUT1}$	+
11			pBSY3G_DaCalB_ $P_{ILV5}$	$P_{ILV5}$	-
12			pBSY3G_PARS1_DaCalB_ $P_{ILV5}$	$P_{ILV5}$	+
13			pBSY3G_DaCalB_ $P_{TPII}$	$P_{TPII}$	-
14			pBSY3G_PARS1_DaCalB_ $P_{TPII}$	$P_{TPII}$	+
15	<i>K. phaffii</i> expression vectors	<i>TPII</i>	pBSY3T_DaCalB_ $P_{AgTEF1}$	$P_{AgTEF1}$	-
16			pBSY3T_PARS1_DaCalB_ $P_{AgTEF1}$	$P_{AgTEF1}$	+
17			pBSY3T_DaCalB_ $P_{GCW14}$	$P_{GCW14}$	-
18			pBSY3T_PARS1_DaCalB_ $P_{GCW14}$	$P_{GCW14}$	+
19			pBSY3T_DaCalB_ $P_{GUT1}$	$P_{GUT1}$	-
20			pBSY3T_PARS1_DaCalB_ $P_{GUT1}$	$P_{GUT1}$	+
21			pBSY3T_DaCalB_ $P_{ILV5}$	$P_{ILV5}$	-
22			pBSY3T_PARS1_DaCalB_ $P_{ILV5}$	$P_{ILV5}$	+
23			pBSY3T_DaCalB_ $P_{TPII}$	$P_{TPII}$	-
24			pBSY3T_PARS1_DaCalB_ $P_{TPII}$	$P_{TPII}$	+

together with the integrative expression strain references (each reference strain was used to inoculate seven wells). The integrative reference strains are average production clones obtained from deep-well plate screening of 41 transformants of linear integrative plasmids (Figure S10 and S11, Supplementary file 1). The average measured *CalB* activities obtained from the deep-well plate screening are summarized in Fig. 1. The ARS plasmid carrying the  $P_{AgTEF1}$  - *GUT1* combination resulted in the highest reporter activity overall being about 83% higher than for the integrative reference. Plasmids with the promoters  $P_{GUT1}$  and  $P_{ILV5}$  driving *GUT1* expression lead to relatively similar volumetric activities, approximately in the range of the  $P_{GUT1}$  - *GUT1* integrative reference strain. The episomal  $P_{TPI1}$  and *GUT1* combination showed a slightly lower reporter activity, comparable to the respective integrative reference. Regarding the *TPI1* marker system, the plasmid with  $P_{ILV5}$  led to the highest *CalB* activities with about 49% surplus relative to the integrative reference. The ARS -  $P_{AgTEF1}$  - *TPI1* combination was the only other to surpass the integrative reference in the *TPI1* system (by 25%). Interestingly, the plasmids with  $P_{GCW14}$ , which is one of the strongest constitutive promoters known in *K. phaffii* [26–28], barely show any *CalB* activity in the *GUT1* marker system and no activity at all in the *TPI1* system. While the  $P_{GCW14}$  - *GUT1* transformants reached comparable cell densities in deep-well plates despite their very small colony size on transformation plates, it was the opposite case for

$P_{GCW14}$  - *TPI1* transformants. Here, the transformants had a regular colony size but these cells showed basically no growth in deep-well plates. Since the expression of the gene *GCW14* has been shown to be very strong when *K. phaffii* is grown on glycerol and the activity of its promoter has been tested on various carbon sources, the regulation of  $P_{GCW14}$  should be suitable for the application tested in this study [26, 27, 29]. So, potentially, expression of the *GUT1* gene, driven by  $P_{GCW14}$  is too strong, perhaps even leading to the accumulation of Gut1p protein in the cell, thereby eliminating the vital necessity of the plasmid carrying the *GUT1* expression cassette. On the other side of the promoter strength spectrum,  $P_{ILV5}$  is said to show only about 15% of  $P_{GAP}$  strength and is therefore the second weakest of the five promoters tested [30]. The *TEF1* promoter from *Ashbya gossypii* is assumed to show higher expression levels than the endogenous equivalent in *K. phaffii*, therefore ranging somewhere near the *GAP* promoter in terms of promoter strength [30, 31]. This strength was the most suitable for this plasmid and reporter setup among the tested promoters. The two native promoters of the tested selection markers,  $P_{GUT1}$  and  $P_{TPI1}$ , were expected to lead to similar reporter activities as the respective integrative references when used with the respective reporter gene. But the episomal  $P_{GUT1}$  - *GUT1* plasmid surpassed the integrative reference in terms of *CalB* activity, while the episomal  $P_{TPI1}$  - *TPI1* plasmid could not maintain the expression levels achieved with the integrative control.  $P_{GUT1}$  is inducible



**Fig. 1** *CalB* reporter activities from transformants of episomal plasmid variants cultivated in micro-scale deep-well plates. Seven transformants of various ARS plasmids with different combinations of selection marker and its promoter, together with 7 colonies of the respective integrative references, were cultivated in 96 deep-well plates containing 250  $\mu$ l BMG1. After 60 h of incubation at 28  $^{\circ}$ C and 320 rpm, 250  $\mu$ l of BMG0.5 were added, followed by 50  $\mu$ l BMG2.5 after 72 and 84 h to allow de-repression of  $P_{DC}$  and consequent reporter gene expression. Cultivations were harvested after 108 h of cultivation by centrifugation and *CalB* activities in the supernatants were evaluated

by glycerol [32–34], but  $P_{TPI1}$  is generally considered to be constitutive, as shown in *Yarrowia lipolytica* [35], but similar work in *K. phaffii* indicates it is affected by the carbon source and shows only <10% the strength of  $P_{GAP}$  when cultivated in complex glycerol media [31].

The specific *CalB* activities (normalized to  $OD_{600}$ ) summarized in Figure S3 (Supplementary file 1) mostly indicate similar ratios between integrative control strains and respective episomal production clones (Fig. 1). For both marker systems, however, the episomal plasmids with  $P_{ILV5}$  driving marker gene expression seem to lead to higher biomass accumulation compared to the others, as the specific activity is reduced in relation to the other plasmid variants.

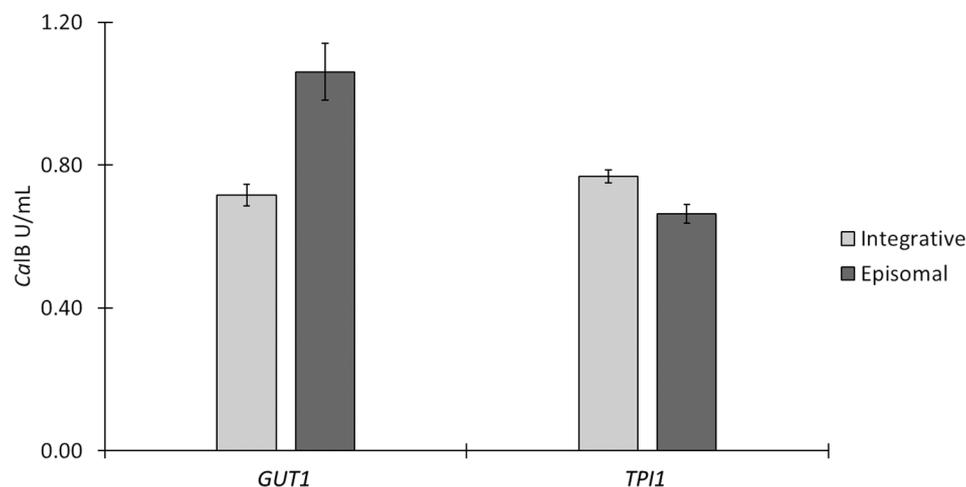
This information described above suggests that a certain level of promoter strength is key for maintaining the episomal plasmids without slowing down metabolism or reducing the selection pressure when trying to produce recombinant protein. The requirements might change with other recombinant proteins of interest as they exert different levels of stress on the host cell and therefore require other gene/transcript dosages for optimal production [36–38].

Interestingly, the standard deviations obtained when screening 7 separate transformants of ARS plasmids were generally low (<9%, except for *GUT1*- $P_{GCW14}$  and *GUT1*- $P_{ILV5}$  plasmids), as they show values similar to the genomically stable integrative reference strains (standard deviations <8%). This indicates their suitability for applications where clonal variability is not desired.

### Shake flask cultivations

One variant of both marker systems was scaled to 250 ml shake flask cultivations. For the *GUT1* marker system, the combination with  $P_{AgTEF1}$  was used for this experiment, as it has been shown to be the most promising in the initial micro scale screening. Even though  $P_{ILV5}$  resulted in slightly higher volumetric *CalB* activities with the *TPI1* system in deep-well plates, the plasmid with  $P_{AgTEF1}$  was preferred, because it showed more reliable specific activities (Figure S3, Supplementary file 1) and it can be assumed that the heterologous sequence of  $P_{AgTEF1}$  would reduce the potential risk of homologous recombination of the marker cassette with the genome compared to the endogenous  $P_{ILV5}$  sequence. Five fresh transformants of each episomal plasmid system were cultivated in separate flasks and each integrative reference strain was cultivated in triplicates in a fed-batch manner as described in the [Materials and Methods](#) section. The results are summarized in Fig. 2.

When employing the *GUT1* episomal system in 250 mL shake flasks with 50 mL media, *CalB* activities surpassed the average control strain with integrated expression cassette by 46% on average. Even though the relative surplus was not as big as in the micro-scale cultivation, the absolute volumetric activities were notably higher. The relative *CalB* activities coming from the episomal *TPI1* system in deep-well plates could not be reproduced in shake flasks as the integrative control led to 14% higher activities compared to the episomal clones. However, similar to the *GUT1* system, the absolute volumetric activities surpassed the deep-well plate cultivation. It was noteworthy that apparently, both episomal systems fell slightly short of expectations, as the surpluses to



**Fig. 2** *CalB* reporter activities from 250 mL shake flask cultivations. Five transformants of ARS plasmids with  $P_{AgTEF1}$  driving the *GUT1* or *TPI1* marker expression were used to directly inoculate baffled 250 mL shake flasks containing 50 mL of BMG1. The respective average, integrative controls were cultivated in triplicates. After 60 h, 72 h and 84 h *GUT1* plasmid cultivations and after 100 h, 112 h and 124 h *TPI1* plasmid cultivations were fed with 500  $\mu$ L of 50% glycerol. After 108 h the *GUT1* plasmid cultivations and after 148 h the *TPI1* plasmid cultivations were harvested and *CalB* activities were measured in the supernatants

the respective references were smaller or even negative. The specific activity levels (normalized to the  $OD_{600}$ ) are summarized in Figure S4 (Supplementary file 1) show a similar pattern as the volumetric activity (Fig. 2), suggesting that differences in biomass are not causing the slight underperformance of the episomal systems. A possible explanation for this outcome could be that only glycerol without any YNB or phosphate buffer was used for feeding, as opposed to BMG media in the deep-well plate experiments, which might have led to a deficit of certain nutrients or an adverse change of pH. Another possibility could be that the productive phase was shorter in the larger shake flask volume compared to the deep-well plate, since also here single colonies were used for direct inoculation and therefore needed longer to deplete the initial glycerol before the  $P_{DC}$  promoter is derepressed [32, 39]. To ensure complete depletion of the initial carbon source (estimated based on final  $OD_{600}$  of the batch phase), the batch phase of the shake flask cultures of *TPII* episomal transformants and the integrative reference strain were extended from 60 to 100 h. Still, the obtained *CalB* activities did not come close to the ones obtained from the *GUT1* system. Plasmid stability and copy numbers have not been assessed for shake flask cultivation, because effects of the cultivation method might have an impact on plasmid stability and copy numbers. Therefore, these values were only assessed for samples of the bioreactor cultivations described later in this study (Table S2, Supplementary file 1), as cultivation parameters can be controlled more strictly in bioreactors. Due to the combined reasons of longer cultivation times and nonetheless lower reporter levels, the *GUT1* based plasmid system was favored for fed-batch bioreactor cultivation tests in this study. While functional and applicable in principle, the *TPII* marker system at least still would need more work for optimization. It could be worthwhile to consider altering the codon usage of the *TPII* gene or testing other stronger, weaker or differently regulated promoters or testing destabilized protein variants of Tpi1p. As the regulation of the  $P_{AgTEFI}$  promoter seems suitable for both marker systems, testing of variants thereof could be also interesting.

Similar to the previously described low standard deviation for deep-well plate cultivations, the standard deviations of 5 single episomal transformants with  $P_{AgTEFI}$  driving marker expression in shake flasks, were also low, ranging from 2% (*TPII* system) to 8% (*GUT1* system) respectively. Compared to the standard deviations of 3 cultures of each of the two genomically stable integrative reference strains (1% for *TPII* and 4% for *GUT1* system) this suggests a high uniformity and low clonal variation in reporter activities among individual transformants of the episomal plasmids.

## Bioreactor cultivations

### The physiological effect of the *GUT1* - deletion on growth capacity

Since the *K. phaffii* platform strain for the *GUT1* carbon source selection marker system has the native *GUT1* gene deleted, the capability of *GUT1* complemented transformants to grow on this substrate was initially characterized in comparison to the unmodified host. The physiologic effect on the producer strains was assessed through three independent glycerol-based batches for each expression system carried out in a 5 L fermenter under controlled conditions, where the yeast was growing at its maximum specific growth rate ( $\mu_{max}$ ), which can be considered as a key parameter to evaluate and understand the growth performance and fitness of the strains [40].

The present study compared 3 different producer strains including: (1) the *gut1* - deletion strain in which plasmid 5 (Table 1) was integrated into the genome (integrative reference strain); (2) the *gut1* - deletion strain, which had been transformed with episomal plasmid 6 (Table 1), and (3) a control strain in which the native *GUT1* remains unmodified. Like the others, this control strain is based on the BSYBG11 platform and contains a single integrated *CalB* expression cassette regulated by the  $P_{DC}$  promoter but the native *GUT1* locus remains unaffected. By evaluating these three cases, the aim was to provide a comprehensive understanding of how this modification might affect the  $\mu_{max}$  and subsequently the capacity of the cells to grow on glycerol as sole carbon source.

As described in previous works the  $\mu_{max}$  during the batch phase can be accurately quantified based on the carbon dioxide evolution rate (CER) on-line data [41]. The  $\mu_{max}$  of the strain with the native *GUT1* gene is  $0.187 \text{ h}^{-1}$ , which is slightly higher than the strain in which the *GUT1* plasmid number 5 was integrated ( $0.179 \text{ h}^{-1}$ ) and that of the strain containing the recombinant episomal *GUT1* plasmid number 6 ( $0.178 \text{ h}^{-1}$ ). The  $\mu_{max}$  was robust along the whole batch cultures. The plots obtained regarding these determinations are included in the Supplementary file 1, Figure S5.

However, since the difference of these results was only in the 5% range, it can be assumed that the capacity to metabolize and grow on glycerol as the sole carbon source has not been significantly affected in the complemented strains. There was no significant difference between integrative and episomal complementation regarding growth on glycerol as sole carbon source and the high fitness of complemented strains was considered to be key for the success of this selection system for scalable production clones.

### Comparative performance of the expression strains in fed-batch cultures

Two biological replicates of fed-batch cultures were performed for each expression system to assess the performance of the episomal *K. phaffii* expression system in bioreactors, and to compare it with the average integrative expression system. The obtained data were also used to determine the key production rates and yields, as well as to assess the plasmid stability of the episomal system. To obtain comparative results, C-limited cultures at a constant specific growth rate of  $0.05 \text{ h}^{-1}$  in which a pseudo-steady state is achieved was selected as a fed-batch strategy. It was based on the implementation of a pre-programmed exponential feeding profile that controls and maintains a constant  $\mu$  by limiting the amount of C fed [42]. The final biomass of around  $80 \text{ g L}^{-1}$  of DCW obtained after 30 h of feeding was very similar in both episomal and integrative expression systems (Table 2).

The  $\mu$  achieved experimentally were in the range of the desired set-point ( $0.05 \text{ h}^{-1}$ ), confirming the successful implementation of the feeding profile and the fed batch strategy. In this sense, there were no significant disparities observed in the specific rates of substrate utilization ( $q_s$ ) or in the biomass-to-substrate yield ( $Y_{X/S}$ ). This confirmed, similar to the previously discussed results of the batch phase, that the growing capacity was not significantly affected by complementation of the *gut1* deficiency using the *GUT1* selection marker either in integrative or episomal plasmids.

The *CalB* production rates and yields of the reporter protein were also determined and compared between the alternative expression systems. Since the *CalB* expression is regulated by the de-repressible  $P_{DC}$  promoter which is repressed during the batch phase (C-excess conditions)

**Table 2** Averaged values of key process parameters obtained by fed-batch cultivations. Nominal and experimental specific growth rate ( $\mu$ ), maximum specific growth rate ( $\mu_{max}$ ), final product titer, final dry cell weight (DCW), substrate consumption rate ( $q_s$ ), specific *CalB* production rate ( $q_p$ ), biomass-to-substrate yield ( $Y_{X/S}$ ), product-to-biomass yield ( $Y_{P/X}$ ) and product-to-substrate yield ( $Y_{P/S}$ ). Results represent the average of two biological replicates.  $\pm$  indicate SD

	Integrative	Episomal	Increase
Nominal $\mu$ ( $\text{h}^{-1}$ )	0.05	0.05	-
Experimental $\mu$ ( $\text{h}^{-1}$ )	$0.0466 \pm 0.0028$	$0.0472 \pm 0.0028$	+1.4%
Final product titer (AU $\cdot \text{L}^{-1}$ )	$4756 \pm 64$	$9906 \pm 463$	+108.3%
Final DCW ( $\text{g L}^{-1}$ )	$77.18 \pm 1.16$	$79.70 \pm 1.24$	+3.3%
$q_s$ ( $\text{g}_S \cdot \text{g}_X^{-1} \cdot \text{h}^{-1}$ )	$0.0898 \pm 0.0067$	$0.0893 \pm 0.0008$	-0.56%
$q_p$ (AU $\text{g}_X^{-1} \cdot \text{h}^{-1}$ )	$3.36 \pm 0.33$	$7.13 \pm 0.54$	+112.4%
$Y_{X/S}$ ( $\text{g}_X \cdot \text{g}_S^{-1}$ )	$0.519 \pm 0.012$	$0.529 \pm 0.027$	+1.9%
$Y_{P/X}$ (AU $\cdot \text{g}_X^{-1}$ )	$72.04 \pm 5.73$	$150.87 \pm 7.07$	+109.4%
$Y_{P/S}$ (AU $\cdot \text{g}_S^{-1}$ )	$45.96 \pm 3.17$	$91.59 \pm 3.11$	+99.3

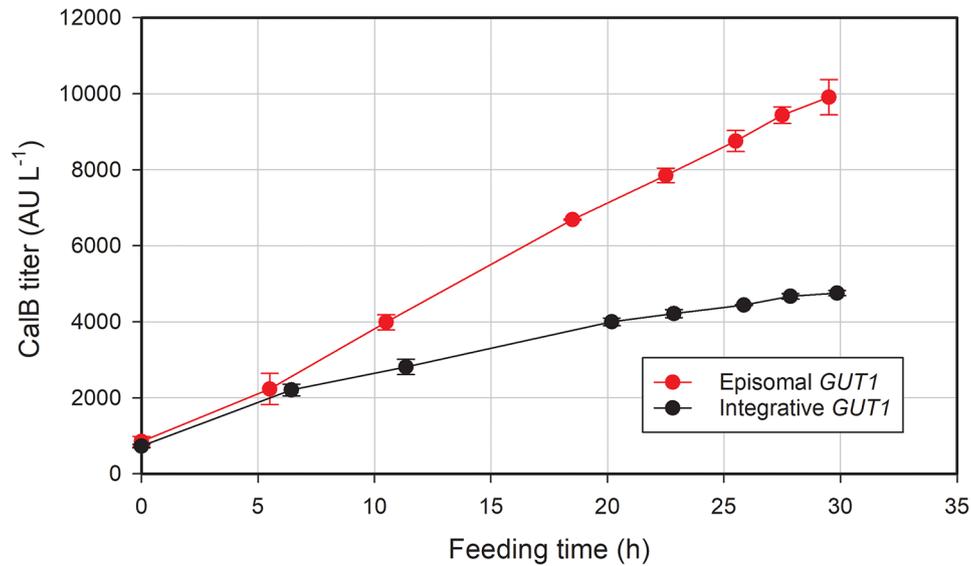
and activated during the carbon limited fed-batch phase, the levels at the end of the batch phase were consistently low in comparison to the final titers. However, after de-repression significant differences in reporter enzyme activity were observed. The episomal system presented a two-fold higher enzymatic activity compared to the average integrative system at the end of the bioprocess. These increases can be seen in Table 2 in terms of titers, specific production rates ( $q_p$ ) and yields (both  $Y_{P/X}$  and  $Y_{P/S}$ ). Additionally, SDS page analysis was performed to verify that *CalB* activities correlate with protein levels in the supernatant (Figure S6, Supplementary file 1).

Interestingly, a consistent and continuous production of *CalB* along the culture for both the integrative and the episomal expression systems was observed (Fig. 3). These constant production profiles are usual for the widely studied genome-integrated systems but were surprising for the episomal system, based on the assumption that the employed ARS plasmids still will get lost to some extent over time. The data of this study indicated that the extrachromosomal plasmids were at least partially maintained in the cells throughout the whole cultivation and production process. For the first time - the stability of an episomal plasmid-based expression strain employing the *GUT1* carbon-source selection system in fed-batch bioreactor cultures was demonstrated to be sufficient to deliver significant amounts of recombinant product from scalable cultivation processes.

### Evaluating the episomal stability of *GUT1* - selected ARS plasmids

Several experiments were performed to investigate the stability of the ARS plasmids with *GUT1* marker in multi day bioreactor cultivations and compared to the strain with integrated expression cassette. First, samples from different cultivation and induction phases (E0 after inoculation, E1 at the end of the batch phase, E4 18.5 h after start of the fed-batch and E8 at the end of the cultivation) in bioreactors were diluted and plated on YPD and selective BMG1 agar plates to detect possible plasmid loss. Second, a comparative expression study was performed with deep-well plate cultivations and *CalB* activity measurements since episomal clones showed different characteristics compare to strains with integrated reporter cassettes. And third, analyses on DNA level had been performed by Southern blot analysis and quantitative PCR.

In the plating studies the number of colonies was determined after 5 days of incubation at room temperature. Cells that still carried the ARS plasmid were able to grow on both media, whereas cells without plasmid only grew on complex, non-selective YPD media. The ratio of cells still carrying the plasmid was calculated by dividing the colony count from selective agar plates (BMG1) by the



**Fig. 3** Time evolution of *CalB* titers during the fed-batch phase of the cultures. Results plotted are the average of two biological replicates

total number of colonies grown on YPD plates. Interestingly, two types of colony sizes, normal and small, were observed on BMG1 plates after plating samples of the episomal system. After inoculation of liquid media only large colonies grew, indicating that the small colonies had been cells which had already lost the plasmid and perhaps grew to small colony size due to the availability of cellular storage compounds, not relying on external carbon sources for the initial cell divisions or impurities of the agar plates enabled such limited growth. Therefore, only large colonies were used for the calculations. In general, a decrease of plasmid-carrying cells (E) over time was observed compared to the integrative construct (I). In sample E0 from the start of the batch phase after inoculation, approximately 22% of the cells carried the plasmid. This number decreased to 12% in sample E1 taken at the start of the Fed-batch, indicating that about 50% of the plasmid containing cells lost the plasmid during the batch where expression of the reporter started. The lowest ratio of plasmid carrying cells was determined to be 7% in E4, showing a reduction of  $\sim 2/3$  compared to the time of inoculation. In sample E8, however, the number of plasmid-carrying cells increased to 9% again. This indicated that during the carbon limited fed-batch cultivation, considering the increase of biomass, the fraction of plasmid containing, and therefore potentially productive cells, remained rather constant or even improved. These data were in line with the activity measurements where the productivity remained constant and did not drop as expected during longer cultivation and induction phases. The samples of the strain with the integrated cassette (I0-I8) did not change over time. These results are summarized in Table S2 (Supplementary file 1). While these findings indicated that the fraction of plasmid containing

cells in the inoculum was already low and also during the batch phase when sufficient carbon source was present, the selection pressure was seemingly low whereas plasmid retention was relatively stable under carbon source limited conditions. Therefore, an improved procedure for the preparation of the inoculum might lead to further process improvements. The non-lethal character of this marker system needs to be considered: Cells without the plasmid do not die, as they would in antibiotic marker applications. They survive for a considerable time while not being able to grow or divide. Their metabolism is reduced and their participation in substrate consumption is assumed to be low, while their reporter production is necessarily zero without plasmid. If these stationary cells are provided with utilizable substrate, like YPD in this experiment, many of them would start growing again and form colonies, which made the subsequently calculated ratio of plasmid carrying cells look small. The high productivity of a cell population with a low fraction of cells carrying the expression cassette might be explained by higher copy numbers/cell for the episomal vectors compared to strains with integrated cassettes or also a better access for transcription at the episomal vector compared to the packed chromosomal expression.

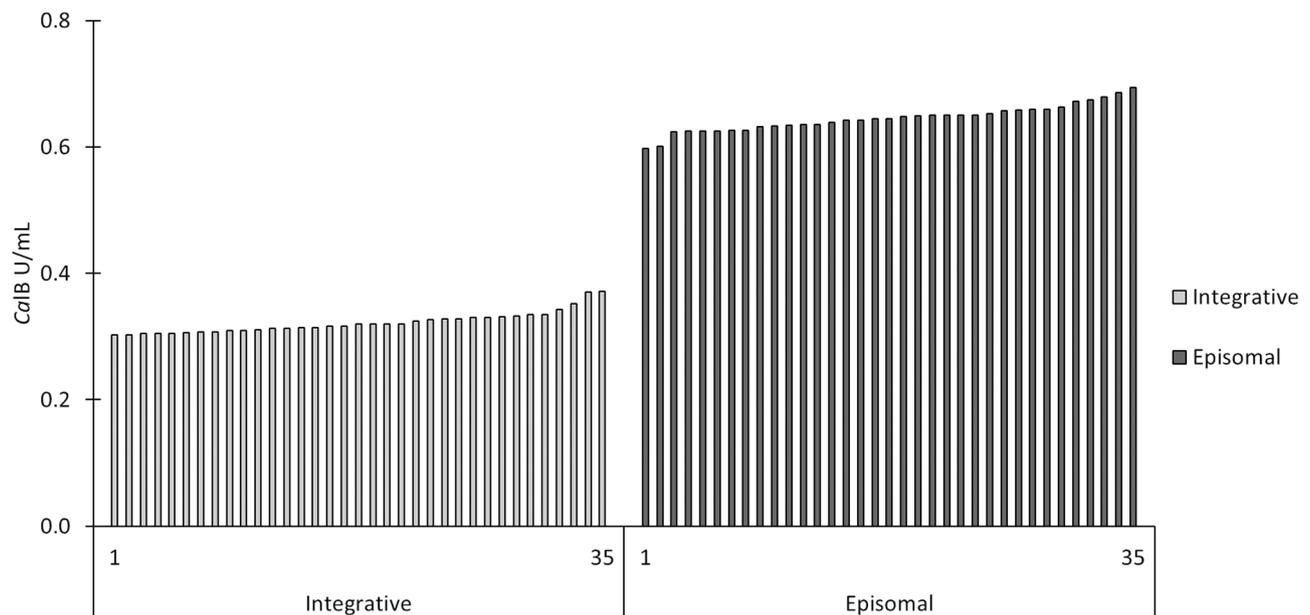
For the second experimental series, 35 single colonies of the plated samples I8 and E8 were inoculated in deep-well plates and cultivated following the same cultivation protocol as already described and reporter enzyme activities in the culture supernatant were determined. Activities in culture supernatants of episomal constructs were typically higher than those measured for strains with integrated cassettes and standard deviations indicating clonal variations or assay variations were low. Interestingly, no measurements for the episomal strains

indicated a potential integration by changing expression characteristics to the typical behaviour of an integrative strain. If such integration event would occur in the late stage of the cultivation such modified phenotype would most likely not be seen by these assays but is also less relevant for the purpose of protein production within a typical cultivation and production period of 3–7 days. For the purpose of possible continuous production processes this would need a more detailed analysis. Therefore, the results summarized in Fig. 4 suggested that there was no significant or early integration event for the episomal plasmids. And, the enzyme titers based on activity measurements were twofold higher ( $201\% \pm 5\%$ ) compared to the *CalB* activity of integrative clones.

Finally, to further investigate the assumption that the *GUT1*-based ARS plasmids are extra-chromosomally stable and do not integrate into the genome on DNA level, also a Southern blot analysis employing a DIG-labeled *CalB* gene as a probe was performed. Genomic DNA was isolated from bioreactor samples E0, E1, E4 and E8, as well as from the integrative reference strain (3 separate preparations) and the empty parental platform strain *K. phaffii* BSY11G1. After detection, the blot (Figure S7, Supplementary file 1) showed bands at approximately 4 and 9.5 Kb that represent the ARS plasmid in the episomal bioreactor samples, which are not visible in the integrative strain or empty strain controls. These bands, however, seem to become weaker as the cultivation progresses, suggesting some plasmid loss in the

cell population over time. This observation is consistent with the results obtained from the plating experiment described above as well as the copy number determination that is described in the following section. When looking at the integrative reference samples, faint but broad bands with sizes larger than 23 Kb, which corresponds to the expression cassette integrated into the genomic DNA fragments of large size caused by shearing forces during genomic DNA isolation, were visible. On the other hand, similar bands are missing in the episomal samples, indicating no or at least a low fraction of cells with integrated expression cassettes. This demonstrated that most of the observed reporter enzyme activity was generated by cells that contained non-integrated episomal expression plasmids.

Although the episomal *GUT1* ARS plasmid expression system is not entirely stable and a noticeable ratio of cells does not carry the plasmid in later phases of the bioreactor cultivation, the reporter activities obtained with the episomal system surprisingly still surpassed the integrative reference significantly. It was known from the literature that the currently available episomal plasmids do not segregate equally to the daughter cells in *K. phaffii* [9, 18], which could mean that, even though a certain proportion of cells might become plasmid-free, the other cells may carry multiple copies of the plasmid. Quantitative PCR was performed to analyze the change of total plasmid DNA in the cell population over time to validate the findings of the other performed analyses. The average



**Fig. 4** Re-screening of individual cells from bioreactor cultivation. Sample 8 from the bioreactor cultivations were plated on BMG1 agar plates to obtain single colonies. A 96 deep-well plate containing 250  $\mu$ l BMG1 medium was inoculated with one single colony per well. After 60 h of incubation at 28  $^{\circ}$ C and 320 rpm, 250  $\mu$ l of BMG0.5 were added, followed by 50  $\mu$ l BMG2.5 after 72 h and 84 h to allow de-repression of  $P_{DC}$  and consequent reporter gene expression. Cultivations were harvested after 108 h of cultivation by centrifugation and *CalB* activities in the supernatants were evaluated. Each bar represents a single well of the deep-well plate

episomal copy number (average of absolute and relative copies) did indeed decrease from approximately  $3.6 \pm 0.4$  to  $1.4 \pm 0.2$  from sample E0 to E4. Interestingly, in sample E8 the copy number increased again to  $2.1 \pm 0.4$ , again confirming the results from plating tests. Accordingly, since the fraction of cells containing plasmids was relatively low already after the batch cultivation this would suggest that cells of the episomal expression strains maintained significantly more than one plasmid copy per cell at the end of the bioreactor cultivation. Considering the average copy number determined is 2.1 per cell at the end of the cultivation and less than 10% of the cells in the population carried a plasmid, each of these cells would have to carry a high number of plasmids, calculated to be approximately 20 copies of plasmid. It was previously shown that *K. phaffii* can maintain high numbers of episomal plasmid copies in a study by Piva et al., where an ARS plasmid expressing eGFP and selected for with Zeocin was determined to be present in approximately 30 copies per cell on average [19].

## Conclusions

In this study the *GUT1* selection marker system for episomal ARS plasmids in *K. phaffii* was refined by employing the *TEF1* promoter of *A. gossypii* for the expression of the glycerol kinase 1 gene that allows growth on glycerol as the sole C-source. This plasmid system enabled recombinant production of lipase *CalB* on bioreactor scale that notably surpasses the average single copy integrative production clone by ~100%, and all performed analyses indicated that the plasmid was maintained as extra-chromosomal vector by a significant fraction of the cell population throughout the cultivation process. For the first time scalable protein production by *K. phaffii* using episomal expression vectors was demonstrated. Additionally, the carbon marker toolbox allowing for efficient and antibiotic free selection was expanded with the *TPII* marker with a *tpi1*-deficient *K. phaffii* platform strain. The high transformation rates and low clonal variations in reporter expression levels among transformants with these two episomal plasmid systems could be beneficial when screening mutagenesis libraries and when comparing the performance of key elements for recombinant protein production such as coding sequence, promoter, secretion signal, etc. Summarizing all advantages and disadvantages of both marker systems, the *GUT1* marker so far showed to be easier to use and therefore preferable. In addition to high transformation rates with episomal vectors the new plasmids enable simple and quick in vivo cloning in *K. phaffii* by simple co-transformation of linear gene constructs with vector backbones having overlapping regions of homology as described previously [10, 14]. Together with the new findings about scalable production this will facilitate the use of *K. phaffii* for a

fast track from gene libraries to bioreactor scale processes for specific gene variants and products. *TPII*, on the other hand, is described as an autoselection system in *S. cerevisiae*, applying selective pressure even in complex media [21, 24]. Therefore, by further improving and investigating the *TPII* marker system with a codon optimized *TPII*, different ARSs, other promoters of different strength and/or regulation, or destabilized or enzymatically less active Tpi1p variants the *TPII* marker system might be improved in future and provide an attractive alternative to the *GUT1* based selection system. Such alternatives might be especially interesting for more demanding continuous cultivation processes.

## Materials and methods

### Chemicals

Restriction enzymes and Phusion HF DNA polymerase (used for standard PCR reactions according to the manual) were ordered from Thermo Fisher Scientific (Vienna, Austria) and single stranded DNA oligonucleotides were purchased from Integrated DNA Technologies (Leuven Belgium). Zeocin™ was obtained from InvivoGen (Toulouse, France), D(+)-biotin from Sigma-Aldrich (Vienna, Austria), Difco yeast nitrogen base w/o amino acids from Becton Dickinson (Schwechat, Austria) and Bacto yeast extract was obtained from (Thermo Fisher Scientific, Vienna, Austria). All other chemicals were purchased from Carl Roth (Karlsruhe, Germany).

### Plasmid construction

The plasmid pBSY3G, kindly provided by Bisy GmbH (Hofstaetten a. d. Raab, Austria), was linearized with *XhoI* and used for Gibson Assembly [43] with a PCR amplicon of the PARS1 sequence (Genbank: M11199.1; Primer 1 & 2), amplified from a CRISPR-Cas9 plasmid assembled by Weninger et al. [44], to form the plasmid pBSY3G\_PARS1 [12]. All subcloning steps were performed in chemo-competent *Escherichia coli* XL1-blue cells. Correct plasmid assembly was verified by Sanger sequencing (Microsynth, Vienna, Austria) after each cloning step. These two plasmids were digested with *SapI* to remove the stuffer fragment. The remaining vector backbones were used for Gibson Assembly with two PCR amplicons: the *Candida antarctica* lipase B gene (*CalB*) from Garrigós-Martínez et al. [28] (Primer 3 & 4) and the  $\alpha$  signal peptide of *S. cerevisiae* taken from the commercially available plasmid pBSYA3S1G (Bisy GmbH) with primers 5 & 6. Then, the resulting pBSY3G\_D $\alpha$ CalB and pBSY3G\_PARS1\_D $\alpha$ CalB were cut with *NcoI* and *SacI* to remove the *GUT1* gene and promoter. The respective 4.1 Kb and 4.3 Kb, fragments were used for Gibson assembly with two inserts: (1) PCR amplicons of the codon-optimized *GUT1* gene from pBSY3G (Primers 7 and 8) or the

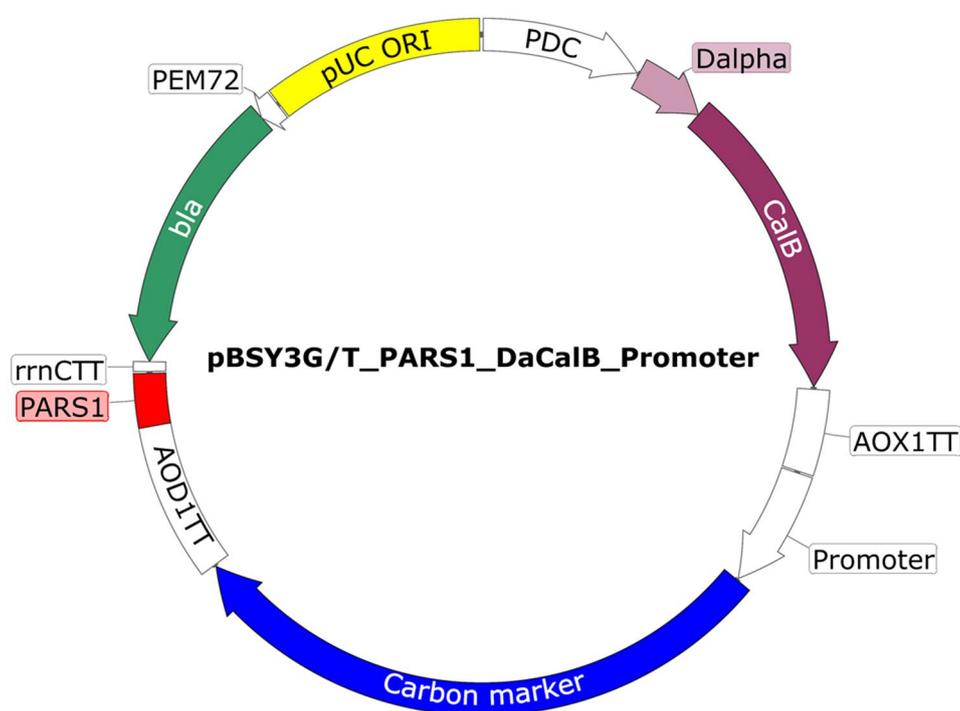
native *TPI1* gene (Genbank protein ID: CAH2449166.1; Primers 9 and 10) and (2) PCR amplicons of promoters  $P_{AgTEF1}$  (from *Ashbya gossypii*) [45] or the endogenous  $P_{GCW14}$  [26],  $P_{GUT1}$ ,  $P_{ILV5}$  [30] or  $P_{TPI1}$  [31] via Gibson Assembly resulting in 20 plasmids total. The promoters were amplified from the respective genomes with primers 11 & 12/13, 14 & 15/16, 17 & 18/19, 20 & 21/22 and 23 & 24/25, respectively. This step was also performed with the  $P_{GUT1}$  to also introduce the novel restriction site between the promoter and marker gene. All primers used in this study are listed in Table S1 (Supplementary file 1). A general map of the ARS plasmids is shown in Fig. 5 and all created plasmids are summarized in Table 1.

### *K. phaffii* strain construction and transformation

The *K. phaffii* strain BSY11G1 ( $\Delta aox1^- \Delta gut1^-$ ), which is unable to grow on glycerol as sole carbon source and shows a Mut<sup>S</sup> phenotype, was used as host for transformation of *GUT1* plasmids with glycerol-based selection. It was kindly provided by Bisy GmbH (Hofstaetten a. d. Raab, Austria) and was made by deletion of the *GUT1* gene of the killer plasmid-free Mut<sup>S</sup> platform strain *K. phaffii* BG11 (Biogrammatix, Carlsbad, CA). This  $\Delta aox1$  deletion strain was transferred to Europe and deposited as BSYBG11 strain by the Bisy GmbH.

For creation of a *tpi1* deficient strain, *K. phaffii* BSYBG11 (Bisy GmbH, Hofstaetten a. d. Raab, Austria)

was transformed with 500 ng of *SmaI* linearized plasmid shown in Figure S8 (Supplementary file 1) in order to facilitate the homologous recombination-based knock out of *TPI1* as described by Cereghino et al. [46]. The transformation was plated on YPEtOH agar plates (2% (w/v) peptone, 1% (w/v) yeast extract, 0.1% (w/v) glucose, 2% (v/v) ethanol, 1.5 (w/v) % agar) containing 100 mg/L Zeocin and was incubated at 28 °C for 7–10 days [47, 48]. Very small colonies were picked and cultivated in 96 deep-well plates (Bel-Art Products, USA) containing YPEtOH media supplemented with 100 mg/L Zeocin for 4 days and then stamped on YPEtOH agar plates with Zeocin and BMG1 [BMG1: 1.34% (w/v) YNB, 1% (w/v) glycerol,  $4 \times 10^{-5}$  % (w/v) biotin, 200 mM potassium phosphate buffer pH 6.0, 1.5% (w/v) agar] agar plates to screen for *TPI1* deficient clones. The knockout of *TPI1* was verified by colony PCR (Phire Plant Direct PCR Master Mix, Thermo Fisher Scientific, Vienna, Austria) (Primers 26 & 27) and subsequent Sanger sequencing (Microsynth, Vienna, Austria). A positive clone was then transformed with an ARS plasmid expressing a flippase gene (Figure S9, Supplementary file 1) for marker recycling. This transformation was performed as the previous one but was plated on YPEtOH agar plates containing 300 mg/L Geneticin. Successful marker recycling was again confirmed by genotyping and the strain was cleared from the ARS plasmid by growing it on non-selective



**Fig. 5** Vector map of a general ARS plasmid carrying a carbon marker cassette. Besides an origin of replication and an Ampicillin resistance cassette for cloning in *E. coli*, the episomal plasmids used in this study harbor the PARS1 sequence for the replication in *K. phaffii*, a reporter cassette with the de-repressible  $P_{DC}$  promoter driving *CalB* expression, whose secretion is facilitated by the D-alpha signal sequence, a deletion variant of the *S. cerevisiae* mating factor alpha signal peptide. Selective pressure in *K. phaffii* is provided by either *GUT1* or *TPI1* genes with one of the 5 tested promoters placed upstream

YPEtOH plates. This *K. phaffii* BSYBG11  $\Delta tpi1$  strain was whole genome sequenced (Illumina, Macrogen, Seoul, South Korea) and is available upon request from BioGrammatics. The Illumina reads were mapped to the *K. phaffii* reference as provided by Sturmberger et al. (Genbank: GCA\_900235035.2) and the data was analyzed according to Schusterbauer et al. [49]. The raw Illumina sequencing reads are available from the ENA database under project PRJEB73854 and relevant analysis results are summarized in the Supplementary file 2. The new *tpi1* deficient strain was used as transformation host with *TPII* based vectors.

*K. phaffii* BSY11G1 and BSYBG11  $\Delta tpi1$  were transformed with 500 ng *SmiI* linearized plasmid variants without ARS sequence (pBSY3G\_ $\Delta\alpha$ CalB\_X and pBSY3T\_ $\Delta\alpha$ CalB\_X, respectively) for integration and 20 ng circular ARS plasmid variants (pBSY3G\_PARS1\_ $\Delta\alpha$ CalB\_X and pBSY3T\_PARS1\_ $\Delta\alpha$ CalB\_X, respectively) similar to the protocol described by Cereghino et al. [46]. The OD<sub>600</sub> of the main cultures were 0.8 and 80  $\mu$ L cells were used for transformations. 920  $\mu$ L of buffered minimal media containing glycerol as sole carbon source (BMG1) were added to the transformations after electroporation and transformations were regenerated for 20 min at 28 °C and 600 rpm on a shaking platform. 20  $\mu$ L of ARS plasmid transformations and 100  $\mu$ L of integrative plasmid transformations were plated on BMG1 agar plates in triplicates for transformation efficiency determination and were incubated for 5 days at 28 °C.

41 transformants of each integrative plasmid were cultivated and screened as described in the following paragraphs to identify average neutral clones for *CalB* production to be used as integrative reference for the respective episomal plasmid transformants. Figures S10 and S11 (Supplementary file 1) show the volumetric *CalB* activity landscapes of the transformants cultivated for the screenings. For each plasmid, clone 21 – the median – was picked and used for further experiments as the respective integrative control. It should be mentioned here, that no transformants of the plasmid 11 with P<sub>LV5</sub> driving *GUT1* expression could be obtained, even after re-sequencing the plasmid and multiple transformation attempts.

#### Small scale cultivations and *CalB* enzyme assay

Freshly grown ARS plasmid transformants and integrative strains were cultivated in 96 deep-well plates (BellArt Products, USA) containing 250  $\mu$ L BMG1 similar to Weiss et al. [50] with some adaptations: After incubating the deep-well plates for 60 h at 28 °C and 320 rpm, the cultivations were fed with 250  $\mu$ L of BMG0.5 [BMG containing 0.5% (w/v) glycerol], as well as 50  $\mu$ L of BMG2.5 [BMG with 2.5% (w/v) Glycerol] after 72 h and 84 h of cultivation in order to allow de-repression of the

carbon-source repressed P<sub>DC</sub> promoter [32, 39], which was used for expression of the *CalB* reporter gene. The P<sub>DC</sub> promoter is a 500 bp fragment of the *K. phaffii* catalase gene *CTAI* (sometimes also annotated as *CAT1*). The OD<sub>600</sub> was measured and the supernatant was harvested after 108 h by centrifugation with 3000 g for 10 min.

For the first upscaling experiments, 250 mL baffled shake flasks containing 50 mL of BMG1 media were inoculated with freshly grown ARS plasmid transformants (1 colony per flask) or a single colony of the integrative reference strain and after 60 h, 72 h and 84 h 250  $\mu$ L of 50% (w/v) Glycerol were added. The initial batch phase was elongated for another 40 h when cultivating transformants of the *TPII* plasmid system due to slow growth rates and to ensure depletion of the initial glycerol supplemented in the media. The cultivations were harvested by centrifugation at 3000 g for 10 min.

Determination of the enzymatic *CalB* activities in the supernatants was performed as described by Krainer et al. [51], but the pH was set at 7 and the kinetics were measured for 3 min in 30 s intervals.

#### Bioreactor cultivations

##### *Inoculum preparation*

The inoculate for bioreactor cultures were grown for approximately 24 h in 1 L baffled shake flasks at 25 °C, 150 rpm, in BMG1 medium to reach an OD<sub>600</sub> of 10. The shake flasks, each filled with 200 mL of culture medium, were initially inoculated with cryo-stocks of the selected producer clones. Subsequently, the cultures were allowed to grow, and cells were harvested via centrifugation, and then resuspended in the batch medium before being introduced into the bioreactor as the inoculum to obtain a starting OD<sub>600</sub> of 1 in the bioreactor. When preparing the inoculum for cultivating the episomal production system, it was ensured to use a fresh single colony transformant to inoculate the shake flask.

##### *Fed-batch cultivations*

The selected operational mode for the bioreactor cultures was a fed-batch. The cultures were performed in duplicates using a 5 L Biostat B fermenter (Sartorius Stedim, Goettingen, Germany), with an initial volume of 2 L. The batch medium contained per liter of distilled water: citric acid 1.8 g, glycerol 40 g, (NH<sub>4</sub>)<sub>2</sub>HPO<sub>4</sub> 12.6 g, MgSO<sub>4</sub>·7H<sub>2</sub>O 0.5 g, KCl 0.9 g, CaCl<sub>2</sub>·2H<sub>2</sub>O 0.02 g, anti-foam Glanapon 2000kz (Bussetti & Co., GmbH, Vienna, Austria) 0.05 ml, biotin 0.4 mg, and 4.6 ml PTM<sub>1</sub> trace salts solution. The PTM<sub>1</sub> solution contained per liter: CuSO<sub>4</sub>·5H<sub>2</sub>O 6.0 g, NaI 0.08 g, MnSO<sub>4</sub>·H<sub>2</sub>O 3.36 g, Na<sub>2</sub>MoO<sub>4</sub>·2H<sub>2</sub>O 0.2 g, H<sub>3</sub>BO<sub>3</sub> 0.02 g, CoCl<sub>2</sub> 0.82 g, ZnCl<sub>2</sub> 20.0 g, FeSO<sub>4</sub>·7H<sub>2</sub>O 65.0 g, and 5.0 ml H<sub>2</sub>SO<sub>4</sub> (95–98%). The feeding fed-batch medium contained per liter of distilled water: glycerol 400 g, MgSO<sub>4</sub>·7H<sub>2</sub>O 6.45 g, KCl

10 g,  $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$  0.35 g, antifoam Glanapon 2000kz 0.2 ml, biotin 0.2 mg, and 1.6 ml  $\text{PTM}_1$  trace salts solution. The specific growth rate ( $\mu$ ) during the fed-batch was fixed at  $0.05 \text{ h}^{-1}$  by a preprogramming exponential feeding rate. During the whole process, Temperature was controlled at  $25 \text{ }^\circ\text{C}$  and pH at 5.0 with  $\text{NH}_4\text{OH}$  15% v/v. Throughout the fermentation, aeration was maintained at 1 vvm. Stirring was varied (600–1200 rpm) in order to keep the dissolved oxygen always above 30%. The procedure is described in detail elsewhere [42].

#### Biomass analyses

Biomass concentration, measured as dry cell weight (DCW), was determined following the method outlined in a previous work [52]. The Relative Standard Deviation (RSD) for this measurement was found to be below 5%.

#### Quantification of the carbon source and byproducts

The quantification of glycerol, and potential fermentation by-products (ethanol, arabitol, and succinate) was determined using High-Performance Liquid Chromatography (HPLC). Details regarding the column and software utilized for this analysis can be found in the work by Garcia-Ortega et al. [53]. The obtained RSD for the HPLC-based analysis was below 1%.

#### Off-gas analyses

The molar fraction of the off-gasses  $\text{CO}_2$  and  $\text{O}_2$  was monitored by a BlueInOne Cell gas analyzer (BlueSens, Herten, Germany) during fed-batch cultivations. The oxygen uptake rate (OUR), carbon dioxide evolution rate (CER), their specific rates ( $q_{\text{O}_2}$  and  $q_{\text{CO}_2}$ ), and respiratory quotient (RQ) were calculated taking the off-gas pressure and humidity measurement into account, as previously described [54]. RSD was less than 5% for all determinations.

#### Key parameters determination

Specific rates and yields were calculated using on-line and off-line data. The calculation was based on the equations derived from material balances. The determination details were described in previous works [55].

In this study the following parameters were determined: the specific growth rate ( $\mu$ ), maximum specific growth rate ( $\mu_{\text{max}}$ ), glycerol consumption specific rate ( $q_s$ ), *Candida antarctica* lipase B (*CalB*) protein specific production rate ( $q_p$ ), carbon dioxide emission rate (CER) and oxygen uptake rate (OUR).

#### Enzymatic activity determination in supernatants of bioreactor cultures

The *CalB* production was quantified by the determination of its enzymatic activity in the sample supernatants.

This determination was performed as described in previous work [56]. RSD was always below 5%.

#### Plasmid copy-number determination

Plasmid copy numbers of the first episomal bioreactor cultivation were determined for bioreactor samples taken from the beginning of the batch phase (E0), at the end of the batch phase (E1), after 18.5 h into the fed-batch (E4) and at the end of the bioreactor cultivation (E8) of the ARS plasmid transformant by real-time PCR (qPCR) with the respective integrative reference strain as reference. The Luna<sup>®</sup> Universal qPCR Master Mix (New England Biolabs, Frankfurt am Main, Germany) was used according to the manufacturer's manual in an Applied Biosystems<sup>™</sup> 7500 Real-Time PCR System (Waltham, Massachusetts, USA). Genomic DNA was isolated according to Hoffman and Winston [57]. qPCR was performed with primers to amplify the reporter gene *CalB* (primers 28 & 29), as well as with primers for amplification of the single copy housekeeping gene *ARG4* (Argininosuccinate lyase gene; primers 30 & 31) as internal reference [58]. In addition to that the genome of the integrative reference strain was Illumina sequenced (Macrogen, Seoul, South Korea). Reads were mapped to the *K. phaffii* reference as provided by Sturmberger et al. (Genbank: GCA\_900235035.2) including the integration vector sequence and bioinformatically analyzed to verify single copy integration via mean coverage as described by Schusterbauer et al. [49]. The raw Illumina sequencing reads are available from the ENA database under project PRJEB73854 and relevant analysis results are summarized in the Supplementary file 2.

#### Southern blot

To detect potential integration events of the ARS plasmid, Southern blot was performed. 20  $\mu\text{g}$  of the gDNA preparations was also used for qPCR and an additional preparation from the empty background strain *K. phaffii* BSY11G1, as well as 10 ng and 200 ng of isolated plasmid were loaded on an 0,75% agarose gel without ethidium bromide (80 V for 2 h). A second gel was prepared with Ethidium bromide and 0.5  $\mu\text{g}$  of the gDNA samples as well as 10 ng and 200 ng of isolated plasmid were loaded on the gel. For detection of the reporter gene *CalB*, a probe was produced by PCR with primers 28 & 32 and the plasmid as a template and was DIG-labeled using the DIG DNA Labeling and Detection Kit (Roche, Mannheim, Germany) following the manufacturer's manual. The blotting and hybridization steps were performed as described by Green and Sambrook [59, 60] in an aqueous hybridization buffer and with a DIG-labeled probe. The probe was detected using the DIG DNA Labeling and Detection Kit.

### Determination of plasmid stability

To calculate the ratio of cells in the bioreactor samples E0, E1, E4 and E8 that still contain the episomal ARS plasmid, approximately  $0.2 \times 10^{-7}$  OD<sub>600</sub> units were plated on YPD (2% peptone, 1% yeast extract, 2% glucose, 1.5% agar) and BMG1 agar plates. The colonies were counted after 5 days at 28 °C and the ratio was calculated by dividing the number of colonies on selective agar plates (BMG1) by the number of colonies on non-selective agar plates (YPD).

### Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12934-024-02368-3>.

Supplementary Material 1

Supplementary Material 2

### Acknowledgements

The authors gratefully acknowledge the financial support by the Christian Doppler Research Association. Special thanks are issued to Anna-Maria Hatzl, Lukas Rieder and the whole research group for valuable discussion and their support in the laboratory as well as Veronika Schusterbauer (bisy GmbH) for her kind help with whole genome analysis. We are also grateful for valuable discussion of all experiments and results with Kjeld Olesen & Sanne Jensen from Novo Nordisk A/S, Knut Madden from BioGrammatics and all other industrial CDlab partners.

### Author contributions

FW: Experimental design, execution of experiments, interpreted results, prepared initial manuscript draft; GRM: Experimental design, execution of experiments, interpreted results, prepared initial manuscript draft; CP: Conceptualized work; FVB: Provided funding, conceptualized work, interpreted results, edited manuscript draft; AG: Managed funding, conceptualized experimental work, provided access to background knowledge, interpreted results and materials, edited manuscript draft; XG-O: Provided funding, conceptualized work, interpreted results, edited manuscript draft.

### Funding

The financial support by the Austrian Federal Ministry for Digital and Economic Affairs, the National Foundation for Research, Technology and Development and the Christian Doppler Research Association (Christian Doppler Laboratory for innovative *Pichia pastoris* host and vector systems) and by the industrial partners Novo Nordisk A/S, BioGrammatics, SeSam Biotech, ZETA and Solar Biotech is gratefully acknowledged.

Open access funding provided by Graz University of Technology.

### Data availability

The datasets and materials used and/or analyzed during the current study are available from the corresponding author on reasonable request. The raw Illumina sequencing reads are available from the ENA database under project number PRJEB73854.

### Declarations

#### Ethics approval and consent to participate

Not applicable.

#### Consent for publication

Not applicable.

#### Competing interests

The authors declare no competing interests.

### Author details

<sup>1</sup>Christian Doppler Laboratory for Innovative *Pichia pastoris* host and vector systems, Institute of Molecular Biotechnology, Graz University of Technology, Graz A-8010, Austria

<sup>2</sup>Christian Doppler Laboratory for Innovative *Pichia pastoris* Host and Vector Systems, Department of Chemical, Biological and Environmental Engineering, Universitat Autònoma de Barcelona, Cerdanyola del Vallès, Bellaterra 08193, Spain

Received: 30 January 2024 / Accepted: 18 March 2024

Published online: 20 April 2024

### References

1. Kurtzman CP. Biotechnological strains of *Komagataella* (*Pichia*) *pastoris* are *Komagataella phaffii* as determined from multigene sequence analysis. *J Ind Microbiol Biotechnol*. 2009;36:1435–8. <https://doi.org/10.1007/s10295-009-0638-4>.
2. Higgins DR, Cregg JM. Introduction to *Pichia pastoris*. In: *Molecular Biotechnology*. 1998. p. 1–15. <https://doi.org/10.1385/0-89603-421-6:1>.
3. Ahmad M, Hirz M, Pichler H, Schwab H. Protein expression in *Pichia pastoris*: recent achievements and perspectives for heterologous protein production. *Appl Microbiol Biotechnol*. 2014;98:5301–17. <https://doi.org/10.1007/s00253-014-5732-5>.
4. Cereghino JL, Cregg JM. Heterologous protein expression in the methylotrophic yeast *Pichia pastoris*. *FEMS Microbiol Rev*. 2000;24:45–66. <https://doi.org/10.1111/j.1574-6976.2000.tb00532.x>.
5. Vogl T, Gebbie L, Palfreyman RW, Speight R. Effect of plasmid design and type of integration event on recombinant protein expression in *Pichia pastoris*. *Appl Environ Microbiol*. 2018;84:2712–29. <https://doi.org/10.1128/AEM.02712-17>.
6. Hohenblum H, Gasser B, Maurer M, Borth N, Mattanovich D. Effects of gene dosage, promoters, and substrates on unfolded protein stress of recombinant *Pichia pastoris*. *Biotechnol Bioeng*. 2004;85:367–75. <https://doi.org/10.1002/bit.10904>.
7. Mattanovich D, Gasser B, Hohenblum H, Sauer M. Stress in recombinant protein producing yeasts. *J Biotechnol*. 2004;113:121–35. <https://doi.org/10.1016/j.jbiotec.2004.04.035>.
8. Geier M, Fauland P, Vogl T, Glieder A. Compact multi-enzyme pathways in *P. pastoris*. *Chem Commun*. 2015;51:1643–6. <https://doi.org/10.1039/C4CC08502G>.
9. Lee CC, Williams TG, Wong DWS, Robertson GH. An episomal expression vector for screening mutant gene libraries in *Pichia pastoris*. *Plasmid*. 2005;54:80–5. <https://doi.org/10.1016/j.plasmid.2004.12.001>.
10. Camattari A, Goh A, Yip LY, Meng Tan AH, Ng SW, Tran A, et al. Characterization of a panARS-based episomal vector in the methylotrophic yeast *Pichia pastoris* for recombinant protein production and synthetic biology applications. *Microb Cell Fact*. 2016;15:1–11. <https://doi.org/10.1186/s12934-016-0540-5>.
11. Näätäsaari L, Mistlberger B, Ruth C, Hajek T, Hartner FS, Glieder A. Deletion of the *Pichia pastoris* KU70 homologue facilitates platform strain generation for gene expression and synthetic biology. *PLoS ONE*. 2012;7. <https://doi.org/10.1371/journal.pone.0039720>.
12. Cregg JM, Barringer KJ, Hessler AY, Madden KR. *Pichia pastoris* as a host system for transformations. *Mol Cell Biol*. 1985;5:3376–85. <https://doi.org/10.1128/mcb.5.12.3376>.
13. Vogl T, Glieder A, WASMAYER R. Novel episomal plasmid vectors, EP3356534B1, Jul. 08, 2020 Accessed: Aug. 22, 2023. [Online]. Available: <https://patents.google.com/patent/EP3356534B1/de?q=EP3356534B1>.
14. Schwarzzhans J-P, Luttermann T, Wibberg D, Winkler A, Hübner W, Huser T, et al. A mitochondrial autonomously replicating sequence from *Pichia pastoris* for Uniform High Level recombinant protein production. *Front Microbiol*. 2017;8 MAY:1–15. <https://doi.org/10.3389/fmicb.2017.00780>.
15. Liachko I, Dunham MJ. An autonomously replicating sequence for use in a wide range of budding yeasts. *FEMS Yeast Res*. 2014;14:364–7. <https://doi.org/10.1111/1567-1364.12123>.
16. Rizvi SMA, Prajapati HK, Ghosh SK. The 2 micron plasmid: a selfish genetic element with an optimized survival strategy within *Saccharomyces cerevisiae*. *Curr Genet*. 2018;64:25–42. <https://doi.org/10.1016/j.plasmid.2013.03.001>.
17. Chan K-M, Liu Y-T, Ma C-H, Jayaram M, Sau S. The 2 micron plasmid of *Saccharomyces cerevisiae*: a miniaturized selfish genome with optimized

- functional competence. *Plasmid*. 2013;70:2–17. <https://doi.org/10.1007/s00294-017-0719-2>.
18. Nakamura Y, Nishi T, Noguchi R, Ito Y, Watanabe T, Nishiyama T, et al. A stable, autonomously replicating plasmid vector containing *Pichia pastoris* centromeric DNA. *Appl Environ Microbiol*. 2018;84. <https://doi.org/10.1128/AEM.02882-17>.
  19. Piva LC, De Marco JL, de Moraes LMP, Reis VCB, Torres FAG. Construction and characterization of centromeric plasmids for *Komagataella phaffii* using a color-based plasmid stability assay. *PLoS ONE*. 2020;15:e0235532. <https://doi.org/10.1371/journal.pone.0235532>.
  20. Uchima CA, Arioka M. Expression and one-step purification of recombinant proteins using an alternative Episomal Vector for the expression of N-Tagged heterologous proteins in *Pichia pastoris*. *Biosci Biotechnol Biochem*. 2012;76:368–71. <https://doi.org/10.1271/bbb.110628>.
  21. Siewiers V. An Overview on Selection Marker Genes for Transformation of *Saccharomyces cerevisiae*. In: Chap. 15: QTL Mapping by Pooled-Segregant Whole-Genome Sequencing in Yeast. 2014. p. 3–15. [https://doi.org/10.1007/978-1-4939-0563-8\\_1](https://doi.org/10.1007/978-1-4939-0563-8_1).
  22. Pronk JT. Auxotrophic yeast strains in fundamental and Applied Research. *Appl Environ Microbiol*. 2002;68:2095–100. <https://doi.org/10.1128/AEM.68.5.2095-2100.2002>.
  23. Kjeldsen T, Balschmidt P, Diers I, Hach M, Kaarsholm NC, Ludvigsen S. Expression of insulin in yeast: the importance of Molecular Adaptation for Secretion and Conversion. *Biotechnol Genet Eng Rev*. 2001;18:89–121. <https://doi.org/10.1080/02648725.2001.10648010>.
  24. Kawasaki GH, Bell L. STABLE DNA CONSTRUCTS, US005871957A, Feb. 16, 1999.
  25. Gunge N. YEAST DNA PLASMIDS. *Annu Rev Microbiol*. 1983;37:253–76. <https://doi.org/10.1146/annurev.mi.37.100183.001345>.
  26. Liang S, Zou C, Lin Y, Zhang X, Ye Y. Identification and characterization of  $P_{GCW14}$ : a novel, strong constitutive promoter of *Pichia pastoris*. *Biotechnol Lett*. 2013;35:1865–71. <https://doi.org/10.1007/s10529-013-1265-8>.
  27. Zhang X, Zhang X, Liang S, Ye Y, Lin Y. Key regulatory elements of a strong constitutive promoter,  $P_{GCW14}$ , from *Pichia pastoris*. *Biotechnol Lett*. 2013;35:2113–9. <https://doi.org/10.1007/s10529-013-1312-5>.
  28. Garrigós-Martínez J, Vuoristo K, Nieto-Taype MA, Tähtiharju J, Uusitalo J, Tukiainen P, et al. Bioprocess performance analysis of novel methanol-independent promoters for recombinant protein production with *Pichia pastoris*. *Microb Cell Fact*. 2021;20:74. <https://doi.org/10.1186/s12934-021-01564-9>.
  29. Liang S, Wang B, Pan L, Ye Y, He M, Han S, et al. Comprehensive structural annotation of *Pichia pastoris* transcriptome and the response to various carbon sources using deep paired-end RNA sequencing. *BMC Genomics*. 2012;13. <https://doi.org/10.1186/1471-2164-13-738>.
  30. Vogl T, Glieder A. Regulation of *Pichia pastoris* promoters and its consequences for protein production. *N Biotechnol*. 2013;30:385–404. <https://doi.org/10.1016/j.nbt.2012.11.010>.
  31. Stadlmayr G, Mecklenbräuer A, Rothmüller M, Maurer M, Sauer M, Mattanovich D, et al. Identification and characterisation of novel *Pichia pastoris* promoters for heterologous protein production. *J Biotechnol*. 2010;150:519–29. <https://doi.org/10.1016/j.jbiotec.2010.09.957>.
  32. Weinhandl K, Winkler M, Glieder A, Camattari A. Carbon source dependent promoters in yeasts. *Microb Cell Fact*. 2014;13:5. <https://doi.org/10.1186/1475-2859-13-5>.
  33. Grauslund M, Lopes JM, Ronnow B. Expression of *GUT1*, which encodes glycerol kinase in *Saccharomyces cerevisiae*, is controlled by the positive regulators *Adr1p*, *Ino2p* and *Ino4p* and the negative regulator *Opi1p* in a carbon source-dependent fashion. *Nucleic Acids Res*. 1999;27:4391–8. <https://doi.org/10.1093/nar/27.22.4391>.
  34. Cregg JM, Tolstorukov I. NOVEL P. PASTORIS PASTORIS PROMOTERS, AND THE USE THEREOF TO DIRECT EXPRESSION OF PROTEINS IN YEAST, PREFERABLY USING A HAPLOID MATING STRATEGY. Privateaccess Point Containing a Sm Card. 2008;1:11–4. <https://patents.google.com/patent/US20080108108A1/en?oq=US20080108108>.
  35. Borkowska M, Białas W, Celińska E. A new set of reference genes for comparative gene expression analyses in *Yarrowia lipolytica*. *FEMS Yeast Res*. 2020;20:1–16. <https://doi.org/10.1093/femsyr/foaa059>.
  36. Tripathi NK, Shrivastava A. Recent developments in Bioprocessing of recombinant proteins: expression hosts and process development. *Front Bioeng Biotechnol*. 2019;7:496566. <https://doi.org/10.3389/fbioe.2019.00420>. <https://www.frontiersin.org/articles/>
  37. Vogl T, Fischer JE, Hyden P, Wasmayer R, Sturmberger L, Glieder A. Orthologous promoters from related methylotrophic yeasts surpass expression of endogenous promoters of *Pichia pastoris*. *AMB Express*. 2020;10:1–10. <https://doi.org/10.1186/s13568-020-00972-1>.
  38. Mellitzer A, Weis R, Glieder A, Flicker K. Expression of lignocellulolytic enzymes in *Pichia pastoris*. *Microb Cell Fact*. 2012;11:61. <https://doi.org/10.1186/1475-2859-11-61>.
  39. Fischer JE, Hatzl A-M, Weninger A, Schmid C, Glieder A. Methanol independent expression by *Pichia pastoris* employing De-repression technologies. *J Visualized Experiments*. 2019;e58589. <https://doi.org/10.3791/58589>.
  40. Looser V, Bruhlmann B, Bumbak F, Stenger C, Costa M, Camattari A, et al. Cultivation strategies to enhance productivity of *Pichia pastoris*: a review. *Biotechnol Adv*. 2015;33:1177–93. <https://doi.org/10.1016/j.biotechadv.2015.05.008>.
  41. Navarrete C, Frost AT, Ramos-Moreno L, Krum MR, Martínez JL. A physiological characterization in controlled bioreactors reveals a novel survival strategy for *Debaryomyces hansenii* at high salinity. *Yeast*. 2021;38:302–15. <https://doi.org/10.1002/yea.3544>.
  42. Garcia-Ortega X, Ferrer P, Montesinos JL, Valero F. Fed-batch operational strategies for recombinant fab production with *Pichia pastoris* using the constitutive *GAP* promoter. *Biochem Eng J*. 2013;79:172–81. <https://doi.org/10.1016/j.bej.2013.07.013>.
  43. Gibson DG, Young L, Chuang R-Y, Venter JC, Hutchison CA, Smith HO. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Methods*. 2009;6:343–5. <https://doi.org/10.1038/nmeth.1318>.
  44. Weninger A, Hatzl A-M, Schmid C, Vogl T, Glieder A. Combinatorial optimization of CRISPR/Cas9 expression enables precision genome engineering in the methylotrophic yeast *Pichia pastoris*. *J Biotechnol*. 2016;235:139–49. <https://doi.org/10.1016/j.jbiotec.2016.03.027>.
  45. Steiner S, Philippson P. Sequence and promoter analysis of the highly expressed *TEF* gene of the filamentous fungus *ashbya gossypii*. *Mol Gen Genet*. 1994;242:263–71. <https://doi.org/10.1007/BF00280415>.
  46. Lin-Cereghino J, Wong WW, Xiong S, Giang W, Luong LT, Vu J, et al. Condensed protocol for competent cell preparation and transformation of the methylotrophic yeast *Pichia pastoris*. *Biotechniques*. 2005;38(44):4648. <https://doi.org/10.2144/05381BM04>.
  47. Compagno C, Boschi F, Ranzi BM. Glycerol production in a triose phosphate isomerase deficient mutant of *Saccharomyces cerevisiae*. *Biotechnol Prog*. 1996;12:591–5. <https://doi.org/10.1021/bp960043c>.
  48. Compagno C, Brambilla L, Capitanio D, Boschi F, Ranzi BM, Porro D. Alterations of the glucose metabolism in a triose phosphate isomerase-negative *Saccharomyces cerevisiae* mutant. *Yeast*. 2001;18:663–70. <https://doi.org/10.1002/yea.715>.
  49. Schusterbauer V, Fischer JE, Gangl S, Schenzle L, Rinnofner C, Geier M, et al. Whole genome sequencing analysis of effects of CRISPR/Cas9 in *Komagataella phaffii*: a budding yeast in distress. *J Fungi*. 2022;8. <https://doi.org/10.3390/jof8100992>.
  50. Weis R, Luiten R, Skranc W, Schwab H, Wubolts M, Glieder A. Reliable high-throughput screening with *Pichia pastoris* by limiting yeast cell death phenomena. *FEMS Yeast Res*. 2004;5:179–89. <https://doi.org/10.1016/j.femsyr.2004.06.016>.
  51. Krainer FW, Dietzsch C, Hajek T, Herwig C, Spadiut O, Glieder A. Recombinant protein expression in *Pichia pastoris* strains with an engineered methanol utilization pathway. *Microb Cell Fact*. 2012;11:22. <https://doi.org/10.1186/1475-2859-11-22>.
  52. Cos O, Serrano A, Montesinos JL, Ferrer P, Cregg JM, Valero F. Combined effect of the methanol utilization (mut) phenotype and gene dosage on recombinant protein production in *Pichia pastoris* fed-batch cultures. *J Biotechnol*. 2005;116:321–35. <https://doi.org/10.1016/j.jbiotec.2004.12.010>.
  53. Garcia-Ortega X, Valero F, Montesinos-Seguí JL. Physiological state as transferable operating criterion to improve recombinant protein production in *Pichia pastoris* through oxygen limitation. *J Chem Technol Biotechnol*. 2017;92:2573–82. <https://doi.org/10.1016/j.nbt.2013.06.007>.
  54. Nieto-Taype MA, Garrigós-Martínez J, Sánchez-Farrando M, Valero F, Garcia-Ortega X, Montesinos-Seguí JL. Rationale-based selection of optimal operating strategies and gene dosage impact on recombinant protein production in *Komagataella phaffii* (*Pichia pastoris*). *Microb Biotechnol*. 2020;13:315–27. <https://doi.org/10.1111/1751-7915.13498>.
  55. Garcia-Ortega X, Adelantado N, Ferrer P, Montesinos JL, Valero F. A step forward to improve recombinant protein production in *Pichia pastoris*: from specific growth rate effect on protein secretion to carbon-starving conditions as advanced strategy. *Process Biochem*. 2016;51:681–91. <https://doi.org/10.1016/j.procbio.2016.08.030>.
  56. Bernat-camps N, Ebner K, Schusterbauer V, Fischer JE, Nieto-taype MA, Valero F, et al. Enabling growth-decoupled *Komagataella Phaffii* recombinant

- protein production based on the methanol-free  $P_{DH}$  promoter. *Front Bioeng Biotechnol.* 2023. <https://doi.org/10.3389/fbioe.2023.1130583>.
57. Hoffman CS, Winston F. A ten-minute DNA preparation from yeast efficiently releases autonomous plasmids for transformation of *Escherichia coli*. *Gene.* 1987;57:267–72. [https://doi.org/10.1016/0378-1119\(87\)90131-4](https://doi.org/10.1016/0378-1119(87)90131-4).
  58. Abad S, Kitz K, Hörmann A, Schreiner U, Hartner FS, Glieder A. Real-time PCR-based determination of gene copy numbers in *Pichia pastoris*. *Biotechnol J.* 2010;5:413–20. <https://doi.org/10.1002/biot.200900233>.
  59. Green MR, Sambrook J. Southern hybridization of radiolabeled probes to nucleic acids immobilized on membranes. *Cold Spring Harb Protoc.* 2021;6:280–9. <https://doi.org/10.1101/pdb.prot100487>.
  60. Green MR, Sambrook J. Southern Blotting. *Cold Spring Harb Protoc.* 2021;2021:269–80. <https://doi.org/10.1101/pdb.prot100495>.

### Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.