A color-based stable multi-copy integrant selection system for *Pichia pastoris* using the attenuated *ADE1* and *ADE2* genes as auxotrophic markers

Min Du, Michael B. Battles† and Juergen H. Nett*

GlycoFi, Inc. (A wholly owned subsidiary of Merck & Co., Inc.); Lebanon, NH USA

†Current affiliation: Adimab, LLC; Lebanon, NH USA

Keywords: *Pichia pastoris*, multi-copy integration, auxotrophic marker, stable integration, *ADE1*, *ADE2*

The methylotropic yeast *Pichia pastoris* has been used for more than two decades to successfully produce a large number of recombinant proteins. Currently, a wide variety of auxotrophic and drug based selection markers are employed to screen for clones expressing the protein of interest. For most proteins an increased copy number of the integrated plasmid results in higher levels of expression, but these multi-copy integrants can be unstable due to the propensity of *P. pastoris* for homologous recombination. Here we describe a multi-copy selection system based on *ade1* and *ade2* auxotrophic parent strains and the respective attenuated markers with truncated promoter regions. We show that for all four proteins we tested, the use of the attenuated markers leads to increased protein expression when compared with selection based on the full strength markers. The fact that the adenine auxotrophic strains grow more slowly than the complemented counterparts essentially ensures the stability of multi-copy integration. At the same time, the accumulation of a red dye in the auxotrophic strains also provides an easy, color-based selection for transformants with multiple copies.

In recent years the budding yeast *Pichia pastoris* has become a popular organism for the expression of heterologous proteins of academic and commercial interest.¹,² We recently showed that it is possible to genetically modify the glycosylation machinery of *P. pastoris* and express heterologous glycoproteins decorated with complex type human glycans.³⁻⁷ However, a need remains for methods and materials to achieve higher cellular productivity in transformed *P. pastoris* cell lines.

Over the years, numerous auxotrophic and dominant selectable markers have been developed⁸⁻¹¹ and used to construct protein expression vectors for various applications. Commonly, a gene of interest is integrated into the *P. pastoris* genome using a plasmid that is either linearized in the auxotrophic marker gene, another homologous genomic region in the plasmid, or in the *AOX1* promoter fragment, and then transformed into the appropriate auxotrophic mutant. Homologous recombination of the free DNA termini then results in single-crossover type integration into these loci. *Most P. pastoris* transformants will contain a single copy of the expression vector, but to obtain transformants that express a high level of the protein of interest it is often desirable to screen for multi-copy integrants. Using expression vectors that contain drug resistance genes like *Kan*° or *Zeo*° as selection markers, it is possible to increase the number of transformants harboring multiple copies of the expression vector by increasing the level of drug used for selection.¹²,¹³ One significant disadvantage of the single-crossover type integration, however, lies in the fact that the multiple integrated copies can collapse back into a single copy by homologous recombination. This has recently been shown by Zhu and coworkers¹⁴ and can be especially problematic during scale-up of the expression reaction during fermentation if the protein of interest is toxic to the cells or the eviction of several copies of expression plasmid results in other growth benefits for the cells.

We therefore developed an expression system based on the slower growing *ade1* and *ade2* auxotrophic strains of *P. pastoris* using expression vectors with *ADE1* and *ADE2* markers with truncated promoter sequences. The fact that multiple copies of the attenuated marker are needed to complement the auxotrophy, essentially ensures a multi-copy integration of the expression vector in the fastest growing transformants. Furthermore, the fact that upon loss of the multiple copies of marker the strain would revert to the slow growth *ade* minus phenotype⁹ ensures that cells that keep multiple copies integrated will be able to outcompete the slower growing ones in the fermenter. An additional benefit of
the system lies in the fact that the *ade1* or the *ade2* genotype leads to the accumulation of a red pigment and results in pink colonies, facilitating identification of the *ade* minus expression host. After transformation with the expression vector, multi-copy integrants can then easily be identified based on their white color and larger colony size.

**Effect of *ADE* Marker Promoter Length on Copy Number and Protein Expression**

To test the effect of the various *ADE* marker promoter truncations on copy number and protein expression, we considered the following assumptions: (1) Since all integration plasmids are integrated into the same genomic locus (i.e., *TRP2*, Fig. 1), it is not expected that a reduction of marker promoter strength will lead to an increased copy number of plasmid integrants per se; (2) if the marker promoter strength drops below a certain threshold it is expected that clones integrating only a single copy of the plasmid will grow at a slower rate than clones integrating multiple copies of the plasmid due to the slow growth phenotype of *ade* minus strains. This should also be concomitant with the appearance of pink color in the low copy clones; (3) a gradual drop in marker promoter strength should therefore lead to decreasing numbers of fast growing white clones and, on a relative basis, increasing numbers of slow growing pink clones; and, (4) in order to eliminate any effect that the expression of a heterologous protein might exert on the growth of transformants, the empty expression plasmids should be tested initially.

We therefore constructed auxotrophic *ade1* strains YGLY563 and YGLY564 (using methods described in File S2) and transformed them with equal amounts (0.2 mg) of integration plasmids pGLY220 to pGLY225 (with marker promoter lengths of 370, 276, 191, 82, 50 and 0 nucleotides respectively; for sequences see File S1) and spread the transformation mixture on minimal media plates. For the general architecture of the expression plasmids please refer to Figure 1. After five days of incubation at 23°C the transformations were assessed for the number of transformants per plate. Surprisingly, integration plasmids pGLY220 to pGLY224 all yielded approximately the same number of colonies. Both yeast strains that had been transformed with pGLY225 however yielded less than 10% of the number of white transformants with a significant number of barely visible, pink transformants in the background (Table 1). It had been anticipated that the plasmids with the promoter truncations would give rise to smaller numbers of colonies as the length of the promoter decreased, with the shortest one, only containing the ORF (designated *ADE1*), yielding none. The results however suggest that the *CYC1* terminator region and the multiple cloning site in front of the marker contain a cryptic promoter activity that allows for a background level of transcription, thereby resulting in levels of *ADE1* gene product that, in multi-copy integrants, are enough to complement the *ade1* auxotrophic phenotype.
We also constructed auxotrophic ade2 strains YGLY1215 and YGLY1216 (using methods described in File S2) and transformed them with equal amounts of integration plasmids pGLY2077 to pGLY2081 (with marker promoter lengths of 126, 82, 51, 13 and 0 nucleotides respectively; for sequences see File S1). In the case of the truncated ADE2 markers however, a gradual reduction in the number of colonies per transformation concomitant with a shorter promoter was observed. As was the case for ADE1, the vector containing only the ADE2 ORF with no native promoter sequence at all (designated ADE2*), yielded less than 10% of the number of white transformants than the construct with the full promoter sequence (Table 1).

In order to test how this anticipated multi-copy integration affected protein expression levels, plasmids expressing human glucocerebrosidase (GBA), single-chain anti-HER2 antibody,\(^1\) human CD40 ectodomain (amino acids 20–192, a gift of R.J. Noelle, see ref.16) or human erythropoietin were transformed into ade1 or ade2 auxotrophic yeast strains (Table 2).

Transformants were grown in 96 well deep well plates as described before,\(^7\) expression was induced using the appropriate carbon source (glucose for GAPDH driven constructs and methanol for AOX1 driven constructs), and protein levels were assessed by protein gel blot or Coomassie gel (Figs. 2 and 3). For most transformations using the ADE1 or ADE2 open reading frames as selection markers, as expected, a very low number of white transformants (5–20) were observed. However the expression level of those clones was significantly higher than clones obtained from transformations using the complete ADE1 or ADE2 gene as markers, which usually gave rise to hundreds of transformants (see Figs. 2 and 3). Especially striking is the amount of protein produced from the clone shown in the lane marked in Figure 2D with a vertical arrow.

We recently used the system described here to create a recombinant human EPO (rhEPO) strain for large scale fermentation (see ref. 17). This strain had been engineered to contain two sets of expression plasmids, one using the Zeocin marker and one using the ADE2* marker. QPCR and deep sequencing allowed us to estimate the copy numbers of each set of expression plasmids that were integrated into the strain. While the first expression plasmid using the Zeocin marker had added approximately 4–5 copies, the second expression plasmid using the ADE2* marker had added another 7–8 copies. PCR analysis of two separate fermentation runs, spanning 40 generations (with induction of protein production between generations 26 and 40), showed that at the end of fermentation no reduction in copy number could be detected (individual clones ranged from 8 to 12 ADE2* marked copies). Furthermore, the fact that the strain could be fermented up to scales of 2000 L without a reduction in rhEPO expression attests to the stable integration of the expression plasmids.

Taken together, our system provides an easy, color based selection of stable multi-copy integrants in P. pastoris. A kit based on the features described here is available commercially from Invitrogen under the name PichiaPink™.

| Yeast Strain | Plasmid designation | Auxotrophic marker | Protein expressed | Promoter | Figure |
|--------------|----------------------|---------------------|-------------------|----------|--------|
| YGLY563      | pJN903               | ADE1                | single chain anti-HER2 antibody | GAPDH    | 2A     |
| YGLY563      | pJN904               | ADE1*               | single chain anti-HER2 antibody | GAPDH    | 2A     |
| YGLY564      | pGLY1084             | ADE1                | Glucocerebrosidase    | GAPDH    | 2B     |
| YGLY564      | pGLY1085             | ADE1*               | Glucocerebrosidase    | GAPDH    | 2B     |
| YGLY563      | pGLY1073             | ADE1                | Human CD40 Ectodomain | AOX1     | 2C     |
| YGLY563      | pGLY1074             | ADE1*               | Human CD40 Ectodomain | AOX1     | 2C     |
| YGLY564      | pGLY1073             | ADE1                | Human CD40 Ectodomain | AOX1     | 2D     |
| YGLY564      | pGLY1074             | ADE1*               | Human CD40 Ectodomain | AOX1     | 2D     |
| YGLY1215     | pGLY2663             | ADE2                | Human EPO            | AOX1     | 3A     |
| YGLY1215     | pGLY2664             | ADE2*               | Human EPO            | AOX1     | 3A     |
| YGLY1216     | pGLY2663             | ADE2                | Human EPO            | AOX1     | 3B     |
| YGLY1216     | pGLY2664             | ADE2*               | Human EPO            | AOX1     | 3B     |

Table 1. Approximate number of white colonies after transformation of yeast strains with plasmids

| Strain   | Plasmid | YGLY563 | YGLY564 |
|----------|---------|---------|---------|
| pGLY220  | 300     | 170     |
| pGLY221  | 300     | 170     |
| pGLY222  | 300     | 170     |
| pGLY223  | 300     | 170     |
| pGLY224  | 300     | 170     |
| pGLY225  | 20      | 3       |
| YGLY1215 | 600     | 600     |
| YGLY1216 | 600     | 500     |
| pGLY2077 | 120     | 80      |
| pGLY2078 | 35      | 40      |
| pGLY2079 | 35      | 25      |

Table 2. Transformation of yeast strains with plasmids expressing heterologous proteins
Materials and Methods

*Escherichia coli* strain DH5α was used for recombinant DNA work. Wild type *P. pastoris* strain NRRL-Y 11430 was used for construction of yeast strains (ATCC #76273). PCR reactions were performed according to supplier recommendations using ExTaq (TaKaRa), Taq Poly (Promega) or Pfu Turbo® (Stratagene). Restriction and modification enzymes were from New England Biolabs. Yeast strains were grown in YPD (1% yeast extract, 2% peptone, 2% dextrose and 1.5% agar) or synthetic defined medium (1.4% yeast nitrogen base, 2% dextrose, 4 × 10^{-7} % biotin and 1.5% agar) supplemented as appropriate. Yeast transformations were performed by electroporation as described in reference 11. Coomassie gels and protein gel blots were performed using 4–20% precast TRIS-SDS gels and the Mini PROTEAN 3 cell (Biorad) under reducing conditions.

*Figure 2.* Protein gel blots and Coomassie gels of the protein produced in *ade1* auxotrophic yeast strains transformed with integration vectors encoding for glucocerebrosidase (GBA), a single-chain anti-HER2 antibody, or human CD40 ectodomain. For each set, multiple clones were randomly picked and analyzed. For details on strains and plasmids please refer to Table 2. Panel A: Protein gel blot showing the expression of a single-chain anti-HER2 antibody using either the full ADE1 gene or only the open reading frame (no promoter) and transcriptional terminator (ADE1*) as selection marker. (B) Protein gel blot showing the expression of GBA using either the full ADE1 gene or ADE1* as selection marker. (C and D) Coomassie gel showing the expression of the human CD40 ectodomain using either the full ADE1 gene or ADE1* as selection marker.
according to the manufacturer's instructions. Primary antibodies for detection were: Goat Anti-Human IgG (Fc) (Pierce, 31413) at 1:10,000 dilution for Herceptin. Anti-human EPO (Santa Cruz Biotechnology, sc7956) at 1:500 dilution. Anti-GBA rabbit polyclonal (Rockland Immunochemicals, Inc., custom made) at 1:500 dilution.

Disclosure of Potential Conflicts of Interest
No potential conflicts of interest to disclose.

Supplementary Material
Supplemental materials can be found at:
http://www.landesbioscience.com/journals/biobugs/article/17936/

Figure 3. Protein gel blots of the protein produced in ade2 auxotrophic yeast strains transformed with integration vectors encoding for human erythropoietin (EPO). For each set, multiple clones were randomly picked and analyzed. For details on strains and plasmids please refer to Table 2. Panels A and B: Protein gel blot showing the expression of EPO using either the full ade2 gene or only the open reading frame (no promoter) and transcriptional terminator (ade2*) as selection marker.

References
1. Cereghino GP, Cereghino JL, Ilgen C, Cregg JM. Production of recombinant proteins in fermenter cultures of the yeast Pichia pastoris. Curr Opin Biotechnol 2002; 13:329-32; PMID:12323354; http://dx.doi.org/10.1016/S0958-1669(02)00330-0
2. Cereghino JL, Cregg JM. Heterologous protein expression in the methylotrophic yeast Pichia pastoris. FEMS Microbiol Rev 2000; 24:45-66; PMID:10640598; http://dx.doi.org/10.1111/j.1574-6976.2000.tb00532.x
3. Choi BK, Bobrowicz P, Davidson RC, Hamilton SR, Kung DH, Li H, et al. Use of combinatorial genetic libraries to humanize N-linked glycosylation in the yeast Pichia pastoris. Proc Natl Acad Sci USA 2003; 100:5022-7; PMID:12702754; http://dx.doi.org/10.1073/pnas.0931263100
4. Hamilton SR, Bobrowicz P, Bobrowicz B, Davidson RC, Li H, Mitchell T, et al. Production of complex human glycoproteins in yeast. Science 2003; 301:1244-6; PMID:12947202; http://dx.doi.org/10.1126/science.1088166
5. Bobrowicz P, Davidson RC, Li H, Pongieter T, Nett JH, Hamilton SR, et al. 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 2004; 14:757-66; PMID:15190003; http://dx.doi.org/10.1093/glycob/cwh104
6. Hamilton SR, Davidson RC, Sethuraman N, Nett JH, Jiang Y, Ross S, et al. Humanization of yeast to produce complex terminally sialylated glycoproteins. Science 2006; 313:1441-3; PMID:16960007; http://dx.doi.org/10.1126/science.1130256
7. Nett JH, Stadheim TA, Li H, Bobrowicz P, Hamilton SR, Davidson RC, et al. A combinatorial genetic library approach to target heterologous glycosylation enzymes to the endoplasmic reticulum or the Golgi apparatus of Pichia pastoris. Yeast 2011; 28:237-52; PMID:21360735; http://dx.doi.org/10.1002/yea.1835
8. Higgins DR, Busser K, Comiskey J, Whittier PS, Purcell TJ, Hoeflter JP. Small vectors for expression based on dominant drug resistance with direct multi-copy selection. Methods Mol Biol 1998; 103:41-53; PMID:9680632
9. Lin Cereghino GP, Lin Cereghino J, Sunga AJ, Johnson MA, Lim M, Glessen MA, et al. New selectable marker/auxotrophic host strain combinations for molecular genetic manipulation of Pichia pastoris. Gene 2001; 263:159-69; PMID:11232554; http://dx.doi.org/10.1016/S0378-1119(00)00576-X
10. Nett JH, Gerngross TU. Cloning and disruption of the PpUra5 gene and construction of a set of integration vectors for the stable genetic modification of Pichia pastoris. Yeast 2003; 20:1279-90; PMID:14618566; http://dx.doi.org/10.1002/yea.1049
11. Nett JH, Hodel N, Rausch S, Wildt S. Cloning and disruption of the Pichia pastoris ARG1, ARG2, ARG3, HIS1, HIS2, HIS5, HIS6 genes and their use as auxotrophic markers. Yeast 2005; 22:295-304; PMID: 15789348; http://dx.doi.org/10.1002/yea.1202
12. Sunga AJ, Tolstorukov I, Cregg JM. Posttransformational vector amplification in the yeast Pichia pastoris. FEMS Yeast Res 2008; 8:870-6; PMID:18637138; http://dx.doi.org/10.1111/j.1567-1364.2008.00410.x
13. Marx H, Mecklenbräuker A, Gasser B, Sauer M, Mattanovich D. Directed gene copy number amplification in Pichia pastoris by vector integration into the ribosomal DNA locus. FEMS Yeast Res 2009; 9:1260-70; PMID:19799649; http://dx.doi.org/10.1111/j.1567-1364.2009.00561.x
14. Zhu T, Guo M, Su C, Qian J, Zhuang Y, Chu J, et al. A systematical investigation on the genetic stability of multi-copy Pichia pastoris strains. Biotechnol Lett 2009; 31:679-84; PMID:19152072; http://dx.doi.org/10.1007/s10529-009-9917-4
15. Zha D, Choi BK, Gerngross T. Single chain antibody with cleavable linker. 2006 US patent application #2006025296.
16. Lu LF, Cook WJ, Lin LL, Noelle RJ. CD40 signaling through a newly identified tumor necrosis factor receptor-associated factor 2 (TRAF2) binding site. J Biol Chem 2003; 278:45414-8; PMID:12960157; http://dx.doi.org/10.1074/jbc.M309601200
17. Nett JH, Gomathinayagam S, Hamilton SR, Gong B, Davidson RC, Du M, et al. Optimization of erythropoietin production with controlled glycosylation - PEGylated erythropoietin produced in glycoengineered Pichia pastoris. J Biotechnol 2011; http://dx.doi.org/10.1016/j.jbiotec.2011.11.002

www.landesbioscience.com Bioengineered Bugs 37