Preprints.org

Article

New Yarrowia lipolytica Recipient Strains Equipped with a Zeta Docking Platform Allow Efficient Integration of a Series of New and Existent Zeta-Based Vectors for Multiple Biotechnological Applications

Clémence Moret and Catherine Madzak [*]

Article

New *Yarrowia lipolytica* Recipient Strains Equipped with a Zeta Docking Platform Allow Efficient Integration of a Series of New and Existent Zeta-Based Vectors for Multiple Biotechnological Applications

Clémence Moret [1,2] and Catherine Madzak [1,*]

[1] Université Paris-Saclay, INRAE, AgroParisTech, UMR SayFood, F-91123 Palaiseau, France
[2] Presently not affiliated
* Correspondence: catherine.madzak@inrae.fr

Abstract: *Yarrowia lipolytica* has a long history of industrial use. With its high secretion capacity, GRAS status and wide range of engineering tools, this yeast has become a powerful host for heterologous protein production as well as a favourite candidate for cellular factory design. Among integration vectors mostly used worldwide for *Y. lipolytica* engineering, zeta-based vectors took advantage of their bordering sequences (LTRs of Ylt1 retrotransposon) to integrate either by homology in Ylt1-bearing strains or non-homologously into any strain devoid of this retrotransposon. As such non-homologous integrations were less efficient and reliable, there was a need for designing zeta-bearing derivatives of widely used *Y. lipolytica* strains. Namely, the W29 wild-type strain and its genetically engineered derivative Po1f strain were equipped with a zeta docking platform in order to allow a more efficient and reliable integration of any zeta-based vector. The presence of this new docking platform allows to compare directly the performances of transformants obtained with different constructs, enabling to perform high-throughput screening or protein engineering. The efficiency of these new Po1z and Po1zΔP strains was demonstrated through the surexpression of two native proteins, a chitinase and a protease. At last, new zeta-based vectors with more powerful promoters were added to this new expression system.

Keywords: white biotechnology; genetic engineering; expression vector; synthetic promoter; non-conventional yeast; cell factory; heterologous expression; *Yarrowia lipolytica*

1. Introduction

Yarrowia lipolytica is a dimorphic and oleaginous yeast from the Saccharomycetes/Hemiascomycetes class, which strains are found in various natural or oil-polluted environments as well as in many foodstuffs, notably in fermented dairy and meat products [1,2]. This industrially-relevant yeast has been used since the 1950s as cattle feed, applied to bioremediation and, more recently, engineered for multiple white biotechnology applications, as extensively reviewed previously [2–9]. As a biosafety class 1 microorganism, *Y. lipolytica* has gained a GRAS (Generally Regarded as Safe) status (FDA, USA) [1]. Most natural isolates are haploid (*MATA* or *MATB*), as are nearly all laboratory strains, although diploidisation can be used as a strategy for assembling interesting strain characteristics [2]. The French strain W29 (aka CLIB 89 and CBS 7504) and a series of its genetically engineered derivatives, Po1d, Po1f, Po1g and Po1h strains, are among the more widely used *Y. lipolytica* strains [2]. Since the first sequencing of a *Y. lipolytica* genome (from E150, aka CLIB 122, a derivative of a crossing between W29 and the American strain

CBS 6124-2, aka CLIB 78), by the Génolevures Consortium in 2004 [10], a total of 27 sequences of *Y. lipolytica* genomes have been made publicly available [11]. These include the sequence of W29 in 2016 [12], and that of its most used derivative, Po1f, in 2014 [13].

A very large array of expression/secretion/surface display vectors and other engineering tools have been developed since a few decades for use in *Y. lipolytica*, and a number of them have been made publicly available through the Addgene website (https://www.addgene.org), as extensively reviewed in 2021 [14]. These tools include notably tunable recombinant promoters, various toolkits for vector construction making use of Golden Gate technology, disruption and marker rescue systems, transposomics tools and various CRISPR-based genome engineering applications [14]. Since then, *Y. lipolytica* technology has continued its exponential development, with the mining of new promoters [15], strategies for hybrid promoter engineering [16], new CRISPR/Cas9-mediated integration toolkits [17,18], the YALIcloneNHEJ and YALIcloneHR modular cloning toolkits for integrating multigene pathways [19,20], an iterative gene integration system [21], an improved CRISPR/Cas9 genome-editing system [22], the design of the searchable YaliFunTome database for transcription factors driving a phenotype of interest [23], and the dissection of W29 carbon metabolism using genome-scale metabolic modelling [24]. Most of these recent works, as well as a lot of their predecessors, make use of W29, or one of its derivative of the Po1 series, as recipient strain.

Despite the present availability of several different modular cloning toolkits for the design of expression vectors for complex engineering of *Y. lipolytica* metabolic pathways, the authors consider that there remains a need for simpler, ready-to-use expression vectors, for establishing proofs of concepts, or for dedicated uses. Such "Yarrowia for dummies" tools could either help beginners to get started with yeast engineering and test rapidly the expression of their genes, or constitute an easy-to-use alternative for projects involving a single gene, like protein engineering. Some vectors developed two decades ago at INRA (presently INRAE), based on an efficient recombinant promoter, are good candidates for such a purpose. Their common promoter, hp4d, is a construct composed of 4 direct copies of an UAS (upstream activating sequence) from *Y. lipolytica XPR2* promoter, linked to a TATA box from *YlLEU2* gene [25]. This strong promoter is mainly active at the beginning of stationary phase, allowing a natural decoupling of growth and expression phases, which made it particularly successful for heterologous expression and one of the prominent tools for *Y. lipolytica* engineering, as reviewed previously [6,14]. Two different expression systems using hp4d were constructed: (i) a series of expression/secretion vectors based on the pBR322 bacterial plasmid, designed to be targeted to a pBR docking platform integrated into the Po1g recipient strain [25]; (ii) a series of expression/secretion/surface display vectors from which the bacterial moiety can be removed to liberate a yeast cassette (expression cassette + selective marker) bordered by zeta sequences (LTRs - long terminal repeats - from Ylt1 retrotransposon) that is used for transformation [26–28]. The pBR-based expression system, characterized by a high transformation efficiency and an efficient targeting to the genomic pBR docking platform by homologous recombination, is commercialized by Yeastern Biotech Co (Taiwan). It has been successfully used for heterologous protein production [29], and is well adapted for protein engineering since the precise targeting of a single copy allows a direct comparison of the characteristics of the transformants obtained with different constructs [30,31]. However, the fact that some bacterial DNA from pBR322 is present in the final construct strongly limits the use of this expression system for a number of applications. In contrast, the zeta-based expression system generates final constructs devoid of any bacterial DNA, which thus retain the GRAS status of *Y. lipolytica*. The zeta-bordered yeast cassette, obtained by *Not*I digestion of any zeta-based vector, carries a URA3 allele as selective marker and can be integrated either by homologous recombination in a Ura⁻ recipient strain bearing the Ylt1 retrotransposon, or by NHEJ (non homologous end joining) in a Ura⁻ recipient strain devoid of Ylt1 (as are most *Y. lipolytica* strains, notably W29 and its Po1 derivatives). In that latter case of random integration, the efficiency of transformation is reduced by approximately an order of magnitude and the insertion of the yeast cassette can possibly disrupt a cell function. In addition, the level of expression of the heterologous gene may be affected (positively or negatively) by nearby genomic sequences [27]. Therefore, a high

number of transformants need to be checked in order to select one which growth is not affected and which is a good producer of the protein of interest. The characteristics of the transformants may thus vary for a same construct, which impair the direct comparison of the effect of different constructs. Despite these limitations, the zeta-based expression system has been transferred to numerous laboratories around the world, through more than hundred MTAs (material transfer agreement), and remains still used nowadays.

The purpose of the present work is to combine the advantages of these two expression sytems: (i) the efficiency and reliability of targeted integration, like with the pBR-based system, and (ii) the large array of vectors and the absence of bacterial DNA, like with the zeta-based system. This can be achieved by constructing some new strains equipped with an integrated zeta docking platform, in order to be able to integrate the *Not*I yeast cassette of (pre-existing or new) zeta-based vectors by homologous recombination, at a precise genomic location. The authors choose to integrate a zeta docking platform into two derivative of W29, the Po1a and the protease-free Po1f strains, in order to generate respectively the Po1z and the Po1zΔP strains. As its parent strain Po1f, Po1zΔP is deleted for both extracellular proteases. As the other GM strains of the Po1 series, both Po1z and Po1zΔP carry a *ura3-302* allele that allows them to metabolize sucrose-based substrates, such as molasses [32].

The newly designed strains have been tested, and the interest of targeting integration of yeast cassettes has been validated, through two examples of (sur)expression of a secreted homologous protein, respectively the *Y. lipolytica* CTS1 chitinase in Po1z, and the AEP protease in Po1zΔP.

2. Materials and Methods

2.1. DNA Synthesis and Engineering

The synthetic cassLEU2t-zeta DNA fragment used for constructing p96zetaLEU, the plasmid used as provider for the zeta docking platform, was provided by Eurogentec. Synthetic DNA fragments used to construct p12-CTS1 and p14-XPR2 vectors (respectively YlCTS1deltaKpnI and matureYlXPR2deltaKpnI) and the one used for promoter engineering (4_UASs_block) were provided by Invitrogen / Life Technologies / Thermo Fisher Scientific. Restriction enzymes and ligase for vector construction were provided by New England Biolabs and used according to the specifications of the supplier.

2.2. Strains and Vectors

All *Y. lipolytica* strains used or newly-designed in this work are described in Table 1. All cultivations of *Y. lipolytica* strains were performed in YPD medium (1% yeast extract, 1% bacto peptone, 1% glucose) at 28°C in shake-flasks at 180 rpm. A derivative of *Escherichia coli* DH5α, NEB 5-alpha (New England Biolabs) was used for constructing and propagating all vectors.

Table 1. Publicly available *Yarrowia lipolytica* strains used in this work.

Strain name * Origin	Genotype Phenotype	Reference number at CIRM-Levures (and other microorganism collections)
W29 French natural isolate [33]	MatA Wild type	CLIB 89 (ATCC 20460, CBS 7504, CICC 1778, NBRC 113670, NRLL Y-3178, VKPM Y-3178)
W29_ura302 GM from W29 [2]	MatA, *ura3-302* *§ Ura⁻, Suc⁺	CLIB 141
Po1a GM from W29_ura302 [2]	MatA, *leu2-270*, *ura3-302* Leu⁻, Ura⁻, Suc⁺	CLIB 140
Po1z GM from Po1a	MatA, *leu2-270::LEU2*-zeta, *ura3-302* Ura⁻, Suc⁺, zeta docking platform	**CLIB 4231**

Po1f GM from Po1a [25]	MatA, *leu2-270, ura3-302, xpr2-322, axp1-2* Leu⁻, Ura⁻, Suc⁺, ◉AEP, ◉AXP	CLIB 724 (ATCC MYA2613, VKPM Y-3155)
Po1z✠P GM from Po1f	MatA, *leu2-270::LEU2*-zeta, *ura3-302,* *xpr2-322, axp1-2* Ura⁻, Suc⁺, ◉AEP, ◉AXP, zeta docking platform	**CLIB 4232**
Po1h GM from Po1f [27]	MatA, *ura3-302, xpr2-322, axp1-2* Ura⁻, Suc⁺, ◉AEP, ◉AXP	CLIB 882
Po1t GM from Po1f [27]	MatA, *leu2-270*, LEU2, *ura3-302::URA3,* *xpr2-322, axp1-2* Suc⁺, ◉AEP, ◉AXP	CLIB 883

*: strains in bold type were designed in this work. [§]: *ura3-302* allele corresponds to a disruption of the *URA3* gene by a heterologous cassette expressing *Saccharomyces cerevisiae SUC2* gene, that confers to the strain the ability to metabolize sucrose [32]. CIRM-Levures website: https://cirm-levures.bio-aware.com/page/Home.

Pre-existing zeta-based vectors were described previously [26–28]; some relevant ones are described in Table 2, together with the newly-designed ones.

Yeast transformations with either *Apa*I fragment from p96zetaLEU or *Not*I yeast cassette from different zeta-based vectors were carried out using the classical lithium acetate treatment and heat shock method described previously [25,27]. Transformants obtained by integration of one of these DNA fragments into the genome of one of the *Y. lipolytica* strains were selected on minimal YNB-N$_{5000}$ medium [25,27]. For each construction, half a dozen independent transformants were picked and checked for normal growth compared to control parent strains. Their growth characteristics were assessed using a BioTek Synergy H1 microplate absorbance reader (Agilent), according to the specifications of the supplier. For the construction of Po1z(ΔP) strains, two transformants were retained for testing the transformation efficiency and one was selected for further use. The comparison of transformation efficiencies between strains was carried out from assays performed in triplicate, using for each case a same amount of yeast cassette DNA, prepared from a same DNA sample of each vector. For the transformation with p12-CTS1 or p14-XPR2, 6 transformants were retained for testing either endo-chitinase or AEP activity.

Table 2. Pre-existing and newly-designed zeta-based vectors publicly available for integration by homologous recombination of a *Not*I yeast cassette in Po1z or Po1zΔP *Yarrowia lipolytica* strains.

Vector *	Promoter Secretion signal (Surface display signal)	Upstream / Downstream cloning sites	Terminator	CIRM-Levures reference number
pINA1311 [26]	hp4d [§] no signal	*Pml*I (blunt) / *Bam*HI, *Kpn*I, *Avr*II	*LIP2* term.	1222032
pINA1312 [26]	hp4d no signal	*Pml*I (blunt) / *Bam*HI, *Kpn*I	*XPR2* term.	1222033
pINA1312- 8UASs aka **pINA2008**	hp8d no signal	*Pml*I (blunt) / *Bam*HI, *Kpn*I	*XPR2* term.	**1223001**
pINA1312- 12UASs aka **pINA2012**	hp12d no signal	*Pml*I (blunt) / *Bam*HI, *Kpn*I	*XPR2* term.	**1223002**

pINA1313 [26]	hp4d *LIP2* prepro	*Xmn*I (in pro) / *Bam*HI, *Kpn*I, *Avr*II	*LIP2* term	1222034	
pINA1314 [26]	hp4d *XPR2* prepro	*Sfi*I (in pro) / *Kpn*I	*XPR2* term.	1222035	
pINA1317 [26]	hp4d *XPR2* pre	*Sfi*I (in pre) / *Avr*II, *Bam*HI, *Kpn*I	*XPR2* term.	1222036	
pINA1317-8UASs aka **pINA7008**	hp8d *XPR2* pre	*Sfi*I (in pre) / *Avr*II, *Bam*HI, *Kpn*I	*XPR2* term.	**1223003**	
pINA1317-12UASs aka **pINA7012**	hp12d *XPR2* pre	*Sfi*I (in pre) / *Avr*II, *Bam*HI, *Kpn*I	*XPR2* term.	**1223004**	
pINA1317-YlCWP110 aka pINA1710, pSD1 [28]	hp4d *XPR2* pre (*CWP1* C-term. 110 bp)	*Sfi*I (in pre) / *Hin*dIII, *Bam*HI, *Aor*13HI**	*XPR2* term.	1222256	

*: vectors in bold type were designed in this work. [§]: hp4d is a recombinant promoter composed of 4 direct copies of the UAS1B sequence, one of the UASs identified in the *Y. lipolytica XPR2* promoter, inserted in tandem upstream from a minimal promoter from *LEU2* gene, corresponding essentially to its TATA box [25]. The UAS1B sequences are bordered by *Sph*I restriction sites, as shown in Fig. S6. In the derived UAS-implemented hp8d and hp12d promoters, the number of UAS1B sequences (also bordered by *Sph*I restriction sites) are respectively of 8 and 12 copies, as shown in Fig. S4 to S6. A general scheme of zeta-based vectors is shown in Fig S1C. The *Sfi*I sites present in either pre or pro modified sequences of *XPR2* secretion signals differ by their 5 central variable nucleotides and are therefore not ligation-compatibles. ******: the previous *Kpn*I site is still present, but located downstream from the surface display signal (*YlCWP1* C-term.), so that insertion of a gene between *Sfi*I and *Kpn*I results in the deletion of the surface display signal: therefore, pINA1710 can be used either as a secretion vector or as a surface display vector, depending of the downstream cloning site used. The full sequence of pINA1314 is shown in Fig. S2 and those of pINA1312 and pINA1317 in Fig. S6. CIRM-Levures website: https://cirm-levures.bio-aware.com/page/Home.

2.3. Enzymatic Activities

Endo-chitinase activity was assessed on 4-Methylumbelliferyl β-D-N,N',N''-triacetylchitotriose (4MU), using a Fluorimetric Chitinase Assay Kit (Sigma-Aldrich) according to the specifications of the supplier, on supernatants of 27h cultures of yeast strains. The fluorescence level of microplate samples were assessed using a SpectraMax M2 spectrofluorometer (Molecular Devices), according to the specifications of the supplier. All assays were performed in triplicate.
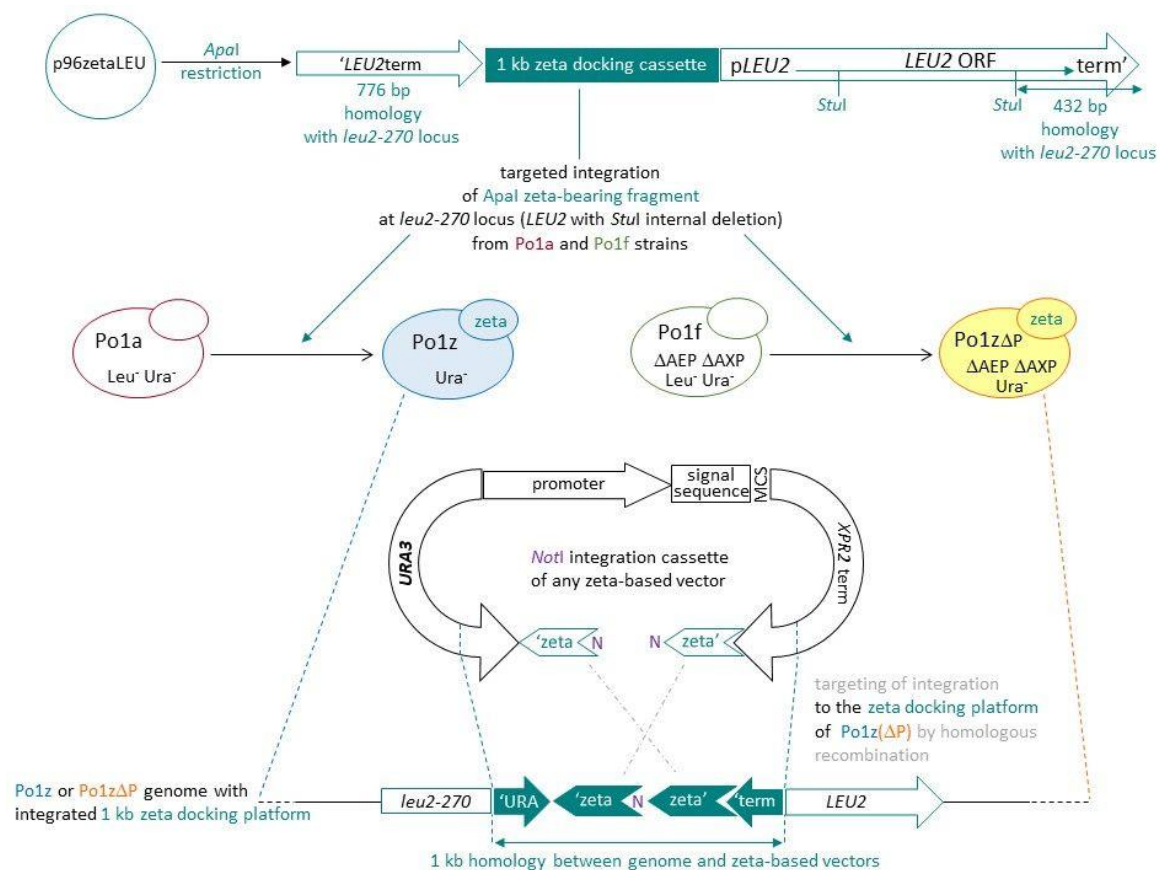
AEP activity was assessed by the measurement of the diameter of halos obtained on skim milk plates (2% Skim Milk, 0.17% Yeast Nitrogen Base, 0.1% Glucose, 1% Agar, 50 mM of Phosphate Buffer at pH 6.8), as previously described [11]. The measurements were performed in triplicate (technical repeats) on three halos for each strain (biological repeats), using the ImageJ software [34] on plate pictures (the plate diameter being used for calibration).

3. Results

3.1. Design of Strains Equipped with a Zeta Docking Platform

In order to construct new *Y. lipolytica* strains equipped with an integrated zeta docking platform, the first step was to design a DNA fragment with a large (1 kb) central region of homology with the different zeta-based vectors and that can be integrated at the *leu2-270* locus of Leu[-] Ura[-] strains to revert their Leucine auxotrophy. As shown in Figure 1 (and with more details, including sequence,

in Figure S1), a synthetic DNA fragment corresponding to the distal part of *LEU2* terminator upstream from a 1 kb zeta cassette (cassLEU2t-zeta) was used to replace the *Nar*I-*Cla*I segment from pINA1296 vector, which placed it upstream from a complete *LEU2* gene sequence. This generated the p96zetaLEU plasmid, from which a 3.3 kb *Apa*I fragment can be extracted and used to transform any Leu⁻ Ura⁻ *Y. lipolytica* strain to provide it with an integrated zeta docking platform. The authors choose to integrate this LEU2-selectable zeta docking platform (i) into a bi-auxotrophic derivative of W29, Po1a, in order to generate the Po1z strain (Ura⁻, Suc⁺) which genetic background is similar to wild-type (only implemented with the ability to metabolize sucrose), and (ii) into the bi-auxotrophic Po1f strain in order to generate the Po1zΔP strain (Ura⁻, Suc⁺, ΔAEP, ΔAXP) that presents a more genetically modified (GM) background (ability to metabolize sucrose and deletion of both extracellular proteases). All the strains used for engineering or as control, as well as the newly-designed ones, are shown in Table 1 with their genetic characteristics.
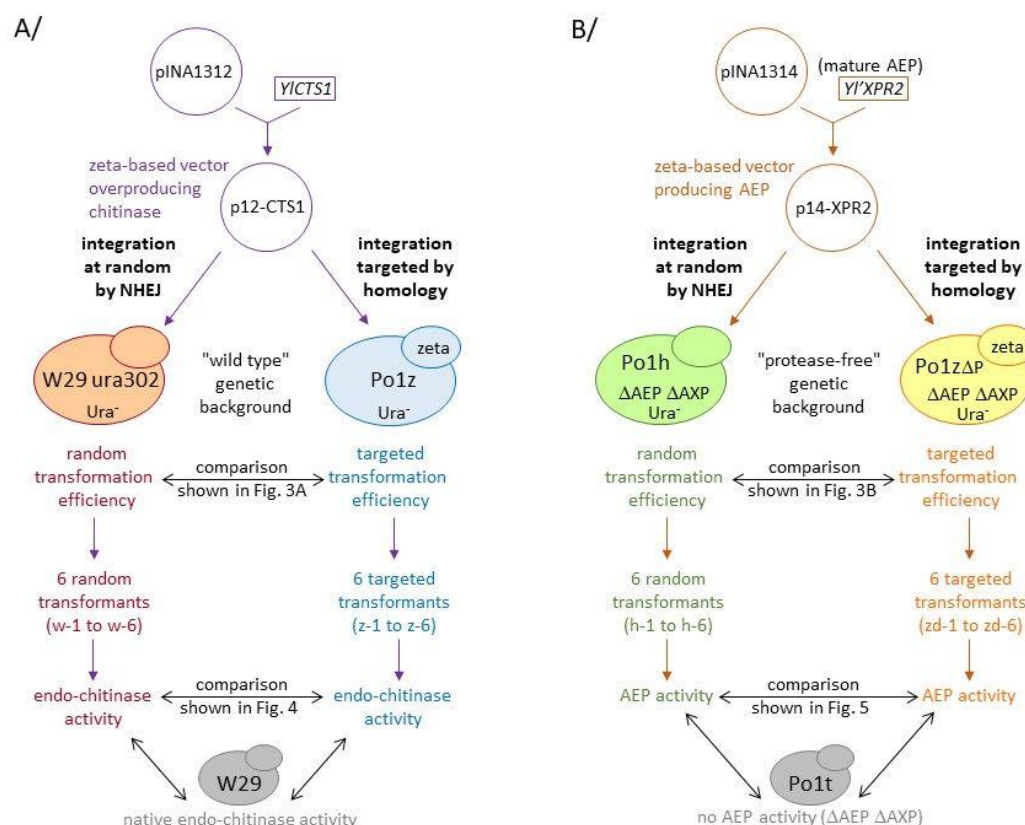

Figure 1. Construction of Po1z and Po1zΔP strains. In this Figure, the ' sign represents a truncation of the upstream or downstream part of a gene or DNA region; ORF stands for Open reading Frame; MCS stands for Multiple Cloning Site, the region of a zeta-based vector in which a heterologous or homologous gene can be inserted for expression/secretion; the purple N represents a *Not*I restriction site or its remnants after cutting. Vector p96zetaLEU (this work) was derived from the *LEU2*-carrying pINA1296 vector [6,25,27] by replacement of the *Nar*I-*Cla*I segment by a synthetic DNA fragment bearing the distal region of *LEU2* terminator upstream from a 1 kb zeta docking cassette (see Figures S1A to D for details and sequence). This new p96zetaLEU vector was submitted to *Apa*I restriction in order to generate the *Apa*I zeta-bearing fragment represented at the top of the Figure. Thanks to its upstream and downstream regions of homology with *Y. lipolytica* genome, this fragment was used to transform Po1a and Po1f strains in order to generate respectively Po1z and Po1zΔP new strains: targeted integration occurred at the *leu2-270* locus and reverted the Leucine auxotrophy of the parent strains. The resulting genomic sequence is represented at the bottom of the Figure: both Po1z and Po1zΔP are equipped with an integrated zeta docking platform that presents a 1 kb homology region with any of the zeta-based

vectors. The bottom part of the Figure represents the targeted integration of such a vector in these new strains (see Figure S1C for more details).
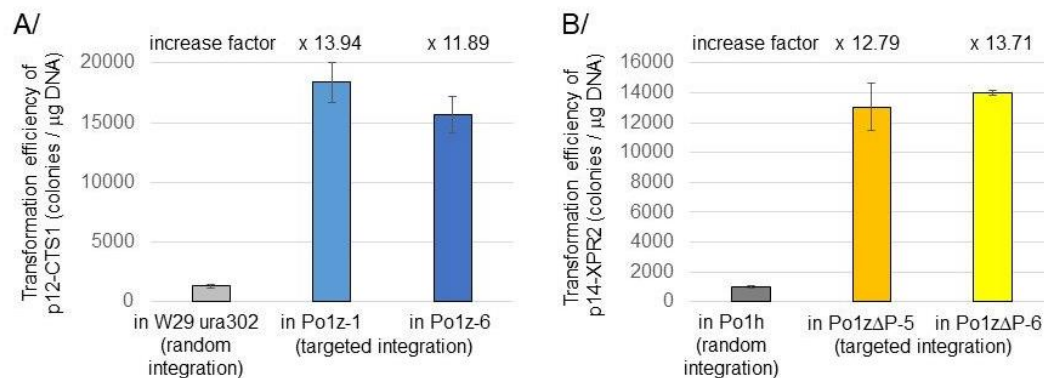
3.2. Validation of New Po1z and Po1zΔP Strains

Thanks to their integrated zeta docking platform that provides a 1 kb homology region with any of the zeta-based vectors (Figures 1 and S1C), both Po1z and Po1zΔP are expected to be transformed more efficiently than the corresponding zeta-devoid W29-ura302 and Po1h strains (targeted versus random integration). A general scheme of the experiments performed to confirm this point and provide examples of the performances of the new zeta-bearing strains is shown in Figure 2. Two secreted homologous enzymes were used to test the new strains, in comparison with their Ylt1-devoid counterparts: the *YlCTS1* chitinase gene was introduced into pINA1312 zeta-based vector and the yeast cassette of the resulting p12-CTS1 plasmid was used to transform either W29_ura302 or Po1z; the *YlXPR2* gene, encoding AEP, was introduced into pINA1314 zeta-based vector and the yeast cassette of the resulting p14-XPR2 plasmid was used to transform either Po1h or Po1zΔP. In both cases, this experiment allowed to compare the transformation efficiency of a same vector either by integration at random (NHEJ) in a Ylt1-devoid strain or by targeted integration (homologous recombination) in a zeta-bearing strain of same genetic background. The results are shown in Figure 3 and demonstrate that the presence of an integrated docking platform generates a 12 to 14 fold increase in transformation efficiency. Two independent transformants from each new strain contruction were tested and the one performing better was selected for further use (respectively Po1z-1 for Po1z strain, and Po1zΔP-6 for Po1zΔP strain).
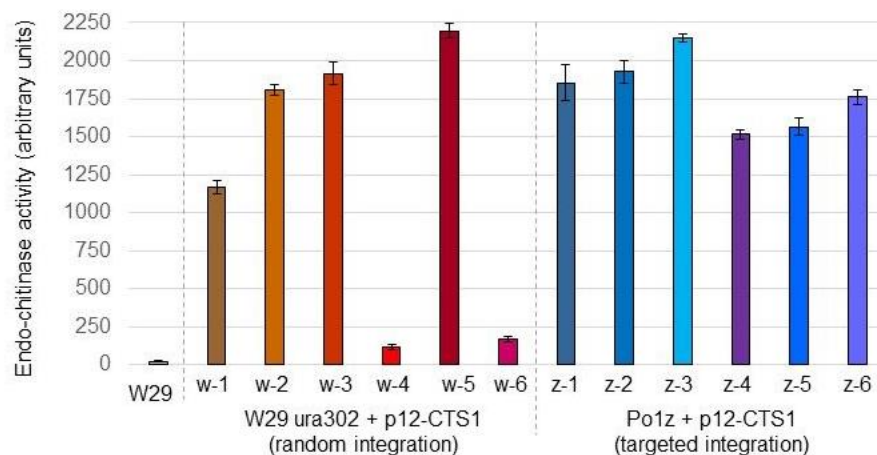

Figure 2. General scheme of the experiments: insertion of a homologous *Y. lipolytica* gene into a zeta-based vector; transformation of yeast, either by random integration into the genome of a Ylt1-devoid strain or by targeted integration into a zeta-bearing strain of same genetic background; comparison of transformation efficiencies and selection of 6 transformants for each category; comparison of enzymatic activities brought by the newly added gene, between all transformants and with a negative control. For the sake of consistency, a

same color code was applied throughout the Figures. **(A)** insertion of the *YlCTS1* chitinase gene into pINA1312 zeta-based expression vector; transformation of either W29_ura302, or the newly designed Po1z strain (cf. Figure 3A); comparison of endo-chitinase activities between 6 transformants of each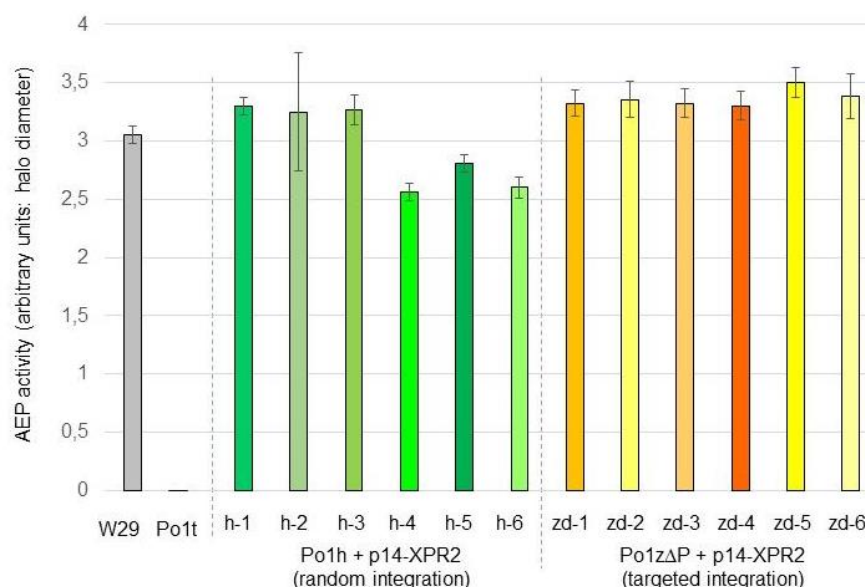 category, and with that from W29 as control (cf. Figure 4). **(B)** insertion of the mature part of the *YlXPR2* gene, encoding extracellular AEP protease, into pINA1314 zeta-based expression/secretion vector (providing the *XPR2* prepro region); transformation of either Po1h or the newly designed Po1zΔP strain (cf. Figure 3B); comparison of AEP activities between 6 transformants of each category, and with that from Po1t as protease-free control (cf. Figure 5).

Figure 3. Comparison of the transformation efficiencies obtained, for a same zeta-based vector, by random integration into a Ylt1-devoid strain or by targeted integration into a derived zeta-bearing strain. In all Figures, the error bars represent the standard deviation (mean value +/- σ). By convention, only effects more than two σ away from mean value are considered statistically significant (confidence level of 95% in the interval of mean value +/- 1.96 x σ). **(A)** Transformation efficiencies obtained with a same amount of yeast cassette DNA from a same sample of p12-CTS1 vector by random integration into W29_ura302 strain or by targeted integration into either Po1z-1 or Po1z-6 strains, two of the transformants selected during the design of the new Po1z strain. On the basis of these results, Po1z-1 was selected for further use as Po1z strain. **(B)** Transformation efficiencies obtained with a same amount of yeast cassette DNA from a same sample of p14-XPR2 vector by random integration into Po1h strain or by targeted integration into either Po1zΔP-5 or Po1zΔP-6 strains, two of the transformants selected during the design of the new Po1zΔP strain. On the basis of these results, Po1zΔP-6 was selected for further use as Po1zΔP strain.

For each transformation experiment (p12-CTS1 in W29_ura302; p12-CTS1 in Po1z; p14-XPR2 in Po1h; p14-XPR2 in Po1zΔP), 6 transformants with unaffected growth (visual checking of colonies and growth tests on Biotek absorbance microplate reader – data not shown) were tested for the newly-engineered enzymatic activity, in comparison with their parent strain (respectively W29 and Po1t), as shown in Figures 4 and 5.

Figure 4. Comparison of endo-chitinase activities obtained after random or targeted integration. Endo-chitinase activities were assessed on 4MU, using a spectofluorometric assay, for W29 (native control) and for 6 transformants obtained with p12-CTS1 vector, either by random integration into W29_ura302 strain (transformants w-1 to -6) or by targeted integration into Po1z strain (transformants z-1 to -6). The error bars represent the standard deviation (mean value +/- σ). Only effects > 2σ away from mean value are considered statistically significant.

Figure 5. Comparison of AEP activities obtained after random or targeted integration. Extracellular protease activities at pH 6.8 (AEP) were assessed by measurement of halos on skim milk plates (cf. examples in Figure S3) for W29 (native control), for Po1t (protease-free control) and for 6 transformants obtained with p14-XPR2 vector, either by random integration into Po1h strain (transformants h-1 to -6) or by targeted integration into Po1zΔP strain (transformants zd-1 to -6). The error bars represent the standard deviation (mean value +/- σ). Only effects more than two σ away from mean value are considered statistically significant (confidence level > 95%).

In the Figure 4, all p12-CTS1 transformants display a successful surexpression of endo-chitinase in comparison to the native *CTS1* level in W29. The maximal activity levels appear similar after either random or targeted integration of p12-CTS1 yeast cassette. However, there is a striking variability among the group of 6 transformants issued from random integration (red-brown hues), in contrast to the pretty similar values observed in the group issued from targeted integration (blue-purple hues). In contrast of the very high activities (in the range of 1500 to more than 2000 arbitrary units) obtained for most transformants, half of the random transformants display much lower activities, especially w-4 (light red) and w-6 (fuchsia) that are underperforming, with activities remaining below 200 arbitrary units. This result point to a major drawback of random integration by NHEJ: the fact that the expression of the newly introduced gene can be negatively affected by regulating elements present in the vicinity of the integration site. In contrast, targeting to a docking platform localized at the genomic URA3 locus, known for its neutrality towards gene expression [6], appears to ensure more reliable and reproducible results.

In the Figure 5, all p14-XPR2 transformants display some AEP activity at levels approximately in the same range than that from the wild-type W29 strain, as expected from the similar strengths of the recombinant hp4d promoter present in the zeta-based vectors and of the native p*XPR2* under induction conditions [25]. However, there is more variability among the group of 6 transformants issued from random integration (green hues), in contrast to the similarly high values observed in the group issued from targeted integration (yellow-orange hues): half of the random transformants (h-4 to -6, in lighter or darker greens) display significantly lower activities. Although less dramatic than

for *CTS1* surexpression, these discrepancies between AEP-expressing random transformants confirm the lesser fiability of NHEJ-mediated integration, compared to targeted one.

3.3. Design of New Zeta-Based Vectors

In order to increase the range of available zeta-based vectors with more powerful expression tools, the authors choose to build new expression vectors with increased promoter strength. For that purpose, the hp4d promoter of pINA1312 and pINA1317 was engineered by iterative addition of a 4 UASs block, in order to generate derivatives with respectively 8 and 12 UASs (promoters hp8d and hp12d), as shown in Figures S4 to S6. The concept of multi-UASs recombinant promoter was initiated with the design of hp4d [25] and generalized later, thanks to the fact that transcription factor availability did not appear to limit the increase in expression obtained with high UAS copy numbers [35,36]. Although expression levels of reporter genes were indeed found to increase continuously with the addition of UASs, up to 32 copies [35,36], it was however shown later, with more practical examples of heterologous protein production, that better yields were attained when using less powerful promoters, such as hp8d [37]. For that purpose, the authors choose to limit the panel of recombinant promoters in zeta-based vectors to hp4d, hp8d and hp12d, as listed in Table 2. In order to engineer further the hp4d promoter present in the expression vector pINA1312 and in the expression/secretion vector pINA1317, a synthetic DNA fragment bearing 4 UASs (4_UASs_block) was inserted upstream from hp4d to create respectively pINA1312-8UASs and pINA1317-8UASs; the operation was repeated to create respectively pINA1312-12UASs and pINA1317-12UASs (Figures S4 and S5). The sequences of the 4_UASs_block DNA fragment, of the parent vectors and of the resulting new ones are shown in Figure S6.

4. Discussion

Through the examples of surexpression of the *Y. lipolytica CTS1* chitinase in Po1z, and of expression/secretion of the AEP protease in Po1zΔP, this work demonstrated the interest of targeting the integration of yeast cassettes at the docking platform of these new zeta-bearing strains: transformation efficiencies were considerably enhanced and the transformants obtained exhibited more reliable and reproducible performances.

Both of these new strains could be transformed with high efficiency by the yeast cassette of any zeta-based vector, through homologous recombination with their zeta docking platform. Depending of the nature of each research project, either the less modified Po1z, or the more modified protease-free Po1zΔP, could be preferred as the recipient strain. The Po1zΔP strain appears more adapted to the production of secreted heterologous proteins, due to the absence of native extracellular proteases likely to degrade the protein of interest.

The array of zeta-based vectors available, already encompassing vectors for intracellular expression, for expression/secretion or for surface display in *Y. lipolytica*, was completed with vectors allowing a more powerful expression through promoter engineering, as shown in Table 2.

Genetic stability of integrated sequences, even those containing direct repeats, is already well documented in *Y. lipolytica* and was demonstrated to be compatible with industrial applications [38]. However, surexpression of some homologous or heterologous proteins was shown to sometime exert a negative influence on the producing strain genetic stability [38]. This potential source of concern should be experimentally addressed by the future users of the expression system, on a case by case basis.

In order to make this user-friendly zeta-targeted expression system freely available to all, the newly designed *Y. lipolytica* strains and vectors were deposited at CIRM-Levures (https://cirm-levures.bio-aware.com). Numerous *Y. lipolytica* strains, including W29 and the whole Po1 series are already available, through MTAs, from the CIRM-Levures catalog. A range of vectors for expression, secretion and surface display in *Y. lipolytica* (including the whole series of zeta-based vectors) are also available; they may not appear yet on the webpage (the addition of a new tab for *Y. lipolytica* engineering tools being still under progress when this paper was submitted) but can nevertheless be requested from CIRM-Levures (contact: cirmlevures@inrae.fr).

The new zeta-targeted expression system developed in this work has been fully validated through the expression examples described. Together with the whole series of zeta-based vectors, the new zeta-bearing

strains constitute a user-friendly and reliable expression system, particularly adapted to basic metabolic engineering or to protein engineering projects.

Abbreviations

The following abbreviations are used in this manuscript:

4MU	4-Methylumbelliferyl β-D-N,N',N''-triacetylchitotriose
AEP	Alkaline Extracellular Protease
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
C-term.	C-terminal extremity of the corresponding protein
DNA	Deoxyribonucleic acid
FDA	Food and Drug Administration
GM	Genetically Modified
GRAS	Generally Regarded as Safe
LTR	Long Terminal Repeats
MCS	Multiple Cloning Site
NHEJ	Non Homologous End Joining
ORF	Open reading Frame

p….	Promoter of the corresponding gene
term.	Terminator
UAS	Upstream Activating Sequence

References

1. Groenewald, M.; Boekhout, T.; Neuvéglise, C.; Gaillardin, C.; van Dijck, P.W.; Wyss, M. *Yarrowia lipolytica*: safety assessment of an oleaginous yeast with a great industrial potential. *Crit. Rev. Microbiol.* **2014**, *40(3)*, 187-206. doi:10.3109/1040841X.2013.770386

2. Madzak, C. *Yarrowia lipolytica* Strains and Their Biotechnological Applications: How Natural Biodiversity and Metabolic Engineering Could Contribute to Cell Factories Improvement. *J. Fungi (Basel)* **2021**, *7(7)*, 548. doi:10.3390/jof7070548

3. Bankar, A.V.; Kumar, A.R.; Zinjarde, S.S. Environmental and industrial applications of *Yarrowia lipolytica*. *Appl. Microbiol. Biotechnol.* **2009**, *84(5)*, 847-865. doi:10.1007/s00253-009-2156-8

4. Liu, H.H.; Ji, X.J.; Huang, H. Biotechnological applications of *Yarrowia lipolytica*: Past, present and future. *Biotechnol. Adv.* **2015**, *33(8)*, 1522-1546. doi:10.1016/j.biotechadv.2015.07.010

5. Zhu, Q.; Jackson, E.N. Metabolic engineering of *Yarrowia lipolytica* for industrial applications. *Curr. Opin. Biotechnol.* **2015**, *36*, 65-72. doi:10.1016/j.copbio.2015.08.010

6. Madzak, C. Engineering *Yarrowia lipolytica* for Use in Biotechnological Applications: A Review of Major Achievements and Recent Innovations. *Mol. Biotechnol.* **2018**, *60(8)*, 621-635. doi:10.1007/s12033-018-0093-4

7. Miller, K.K.; Alper, H.S. Yarrowia lipolytica: more than an oleaginous workhorse. *Appl. Microbiol. Biotechnol.* **2019**, *103(23-24)*, 9251-9262. doi:10.1007/s00253-019-10200-x

8. Park, Y.K.; Ledesma-Amaro, R. What makes *Yarrowia lipolytica* well suited for industry? *Trends Biotechnol.* **2023**, *41(2)*, 242-254. doi:10.1016/j.tibtech.2022.07.006

9. Sun, M.L.; Gao, X.; Lin, L.; Yang, J.; Ledesma-Amaro, R.; Ji, X.J. Building *Yarrowia lipolytica* Cell Factories for Advanced Biomanufacturing: Challenges and Solutions. *J. Agric. Food Chem.* **2024**, *72(1)*, 94-107. doi:10.1021/acs.jafc.3c07889

10. Dujon, B.; Sherman, D.; Fischer, G.; Durrens, P.; Casaregola, S.; Lafontaine, I.; De Montigny, J.; Marck, C.; Neuveglise, C.; Talla, E. *et al*. Genome evolution in yeasts. *Nature* **2004**, *430*, 35–44. doi:10.1038/nature02579

11. Bigey, F.; Pasteur, E.; Połomska, X.; Thomas, S.; Crutz-Le Coq, A.M.; Devillers, H.; Neuvéglise, C. Insights into the Genomic and Phenotypic Landscape of the Oleaginous Yeast *Yarrowia lipolytica*. *J. Fungi (Basel)*. **2023**, *9(1)*, 76. doi:10.3390/jof9010076

12. Magnan, C.; Yu, J.; Chang, I.; Jahn, E.; Kanomata, Y.; Wu, J.; Zeller, M.; Oakes, M.; Baldi, P.; Sandmeyer, S. Sequence Assembly of *Yarrowia lipolytica* Strain W29/CLIB89 Shows Transposable Element Diversity. *PLoS ONE* **2016**, *11*, e0162363. doi:10.1371/journal.pone.0162363

13. Liu, L.; Alper, H.S. Draft Genome Sequence of the Oleaginous Yeast *Yarrowia lipolytica* PO1f, a Commonly Used Metabolic Engineering Host. *Genome Announc.* **2014**, *2*, e00652-14. doi:10.1128/genomeA.00652-14

14. Madzak, C. *Yarrowia lipolytica* engineering as a source of microbial cell factories. In *Microbial Cell Factories Engineering for Production of Biomolecules*; Singh, V., Ed.; Academic Press: Cambridge, MA, USA, 2021; pp. 345–380.

15. Georgiadis, I.; Tsiligkaki, C.; Patavou, V.; Orfanidou, M.; Tsoureki, A.; Andreadelli, A.; Theodosiou, E.; Makris, A.M. Identification and Construction of Strong Promoters in *Yarrowia lipolytica* Suitable for Glycerol-Based Bioprocesses. *Microorganisms*. **2023**, *11(5)*, 1152. doi:10.3390/microorganisms11051152

16. Zhao, Y.; Liu, S.; Lu, Z.; Zhao, B.; Wang, S.; Zhang, C.; Xiao, D.; Foo, J.L.; Yu, A. Hybrid promoter engineering strategies in *Yarrowia lipolytica*: isoamyl alcohol production as a test study. *Biotechnol. Biofuels*. **2021**, *14(1)*, 149. doi:10.1186/s13068-021-02002-z

17. Cui, Z.; Zheng, H.; Zhang, J.; Jiang, Z.; Zhu, Z.; Liu, X.; Qi, Q.; Hou, J. A CRISPR/Cas9-Mediated, Homology-Independent Tool Developed for Targeted Genome Integration in *Yarrowia lipolytica*. *Appl. Environ. Microbiol.* **2021**, *87(6)*, e02666-20. doi:10.1128/AEM.02666-20

18. Liu, X.; Cui, Z.; Su, T.; Lu, X.; Hou, J.; Qi, Q. Identification of genome integration sites for developing a CRISPR-based gene expression toolkit in *Yarrowia lipolytic*a. *Microb. Biotechnol*. **2022**, *15(8)*, 2223-2234. doi:10.1111/1751-7915.14060

19. Li, Y.W.; Yang, C.L.; Shen, Q.; Peng, Q.Q.; Guo, Q.; Nie, Z.K.; Sun, X.M.; Shi, T.Q.; Ji, X.J.; Huang, H. YALIcloneNHEJ: An Efficient Modular Cloning Toolkit for NHEJ Integration of Multigene Pathway and Terpenoid Production in *Yarrowia lipolytica*. *Front. Bioeng. Biotechnol*. **2022**, *9*, 816980. doi:10.3389/fbioe.2021.816980

20. Shen, Q.; Yan, F.; Li, Y.W.; Wang, J.; Ji, J.; Yan, W.X.; He, D.C.; Song, P.; Shi, T.Q. Expansion of YALIcloneHR toolkit for *Yarrowia lipolytica* combined with Golden Gate and CRISPR technology. *Biotechnol. Lett*. **2024**, *46(1)*, 37-46. doi:10.1007/s10529-023-03444-1

21. Luo, Z.; Shi, J.T.; Chen, X.L.; Chen, J.; Liu, F.; Wei, L.J.; Hua, Q. Iterative gene integration mediated by 26S rDNA and non-homologous end joining for the efficient production of lycopene in *Yarrowia lipolytica*. *Bioresour. Bioprocess*. **2023**, *10*, 83. doi:10.1186/s40643-023-00697-6

22. Zhang, B.; Cao, J. Improving and Streamlining Gene Editing in *Yarrowia lipolytica* via Integration of Engineered Cas9 Protein. *J. Fungi (Basel)*. **2024**, *10(1)*, 63. doi:10.3390/jof10010063

23. Gorczyca, M.; Białas, W.; Nicaud, J.M.; Celińska, E. 'Mother(Nature) knows best' - hijacking nature-designed transcriptional programs for enhancing stress resistance and protein production in *Yarrowia lipolytica*; presentation of YaliFunTome database. *Microb. Cell Fact*. **2024**, *23(1)*, 26. doi:10.1186/s12934-023-02285-x

24. Guo, Y.; Su, L.; Liu, Q.; Zhu, Y.; Dai, Z.; Wang, Q. Dissecting carbon metabolism of *Yarrowia lipolytica* type strain W29 using genome-scale metabolic modelling. *Comput. Struct. Biotechnol. J*. **2022**, *20*, 2503-2511. doi:10.1016/j.csbj.2022.05.018

25. Madzak, C.; Tréton, B.; Blanchin-Roland, S. Strong hybrid promoters and integrative expression/secretion vectors for quasiconstitutive expression of heterologous proteins in the yeast *Yarrowia lipolytica*. *J. Mol. Microbiol. Biotechnol.* **2000**, *2*, 207–216.

26. Nicaud, J.M.; Madzak, C.; van den Broek, P.; Gysler, C.; Duboc, P.; Niederberger, P.; Gaillardin, C. Protein expression and secretion in the yeast *Yarrowia lipolytica*. *FEMS Yeast Res*. **2002**, *2*, 371–379. doi:10.1016/S1567-1356(02)00082-X

27. Madzak, C. New tools for heterologous protein production in the yeast *Yarrowia lipolytica*. In *Recent Research Developments in Microbiology*; Pandalai, S.G., Ed.; Research Signpost: Trivandrum, India, 2003; Volume 7, pp. 453–479.

28. Yue, L.; Chi, Z.; Wang, L.; Liu, J.; Madzak, C.; Li, J.; Wang, X. Construction of a new plasmid for surface display on cells of *Yarrowia lipolytica*. *J. Microbiol. Methods* **2008**, *72*, 116–123. doi:10.1016/j.mimet.2007.11.011

29. Kopecný, D.; Pethe, C.; Sebela, M.; Houba-Hérin, N.; Madzak, C.; Majira, A.; Laloue, M. High-level expression and characterization of *Zea mays* cytokinin oxidase/dehydrogenase in *Yarrowia lipolytica*. *Biochimie* **2005**, *87*, 1011–1022. doi:10.1016/j.biochi.2005.04.006

30. Madzak, C.; Mimmi, M.C.; Caminade, E.; Brault, A.; Baumberger, S.; Briozzo, P.; Mougin, C.; Jolivalt, C. Shifting the optimal pH of activity for a laccase from the fungus *Trametes versicolo*r by structure-based mutagenesis. *Protein Eng. Des. Sel.* **2006**, *9*, 77–84. doi:10.1093/protein/gzj004

31. Galli, C.; Gentili, P.; Jolivalt, C.; Madzak, C.; Vadalà, R. How is the reactivity of laccase affected by single-point mutations? Engineering laccase for improved activity towards sterically demanding substrates. *Appl. Microbiol. Biotechnol.* **2011**, *91*, 123–131. doi:10.1007/s00253-011-3240-4

32. Nicaud, J.M.; Fabre, E.; Gaillardin, C. Expression of invertase activity in *Yarrowia lipolytica* and its use as a selective marker. *Curr. Genet*. **1989**, *16*, 253–260. doi: 10.1007/BF00422111

33. Gaillardin, C.M.; Charoy, V.; Heslot, H. A study of copulation, sporulation and meiotic segregation in *Candida lipolytica*. *Arch. Mikrobiol*. **1973**, *92*, 69–83. doi: 10.1007/BF00409513

34. Schneider, C.A.; Rasband, W.S.; Eliceiri, K.W. NIH Image to ImageJ: 25 years of image analysis. *Nat. Meth.* **2012**, *9(7)*, 671–675. doi: 10.1038/nmeth.2089

35. Blazeck, J.; Liu, L.; Redden, H.; Alper, H. Tuning gene expression in *Yarrowia lipolytica* by a hybrid promoter approach. *Appl. Environ. Microbiol.* **2011**, *77(22)*, 7905–7914. doi: 10.1128/AEM.05763-11

36. Blazeck, J.; Reed, B.; Garg, R.; Gerstner, R.; Pan, A.; Agarwala, V.; Alper, H.S. Generalizing a hybrid synthetic promoter approach in *Yarrowia lipolytica*. *Appl. Microbiol. Biotechnol*. **2013**, *97(7)*, 3037–3052. doi: 10.1002/biot.201200120

37. Dulermo, R.; Brunel, F.; Dulermo, T.; Ledesma-Amaro, R.; Vion, J.; Trassaert, M.; Thomas, S.; Nicaud, J.M.; Leplat, C. Using a vector pool containing variable-strength promoters to optimize protein production in *Yarrowia lipolytica*. *Microb. Cell Fact*. **2017**, *16(1)*, 31. doi: 10.1186/s12934-017-0647-3

38. Le Dall, M.T.; Nicaud, J.M.; Gaillardin, C. Multiple-copy integration in the yeast *Yarrowia lipolytica*. *Curr Genet*. **1994**, *26(1)*, 38-44. doi: 10.1007/BF00326302