



Recombination-stable multimeric green fluorescent protein for characterization of weak promoter outputs in *Saccharomyces cerevisiae*

Rugbjerg, Peter; Knuf, Christoph; Förster, Jochen; Sommer, Morten Otto Alexander

Published in:
FEMS Yeast Research

Link to article, DOI:
[10.1093/femsyr/fov085](https://doi.org/10.1093/femsyr/fov085)

Publication date:
2015

Document Version
Peer reviewed version

[Link back to DTU Orbit](#)

Citation (APA):

Rugbjerg, P., Knuf, C., Förster, J., & Sommer, M. O. A. (2015). Recombination-stable multimeric green fluorescent protein for characterization of weak promoter outputs in *Saccharomyces cerevisiae*. *FEMS Yeast Research*, 15(8), [fov085]. <https://doi.org/10.1093/femsyr/fov085>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

1

2

3

4 **Recombination-stable multimeric green fluorescent protein**

5 **for characterization of weak promoter outputs in**

6 *Saccharomyces cerevisiae*

7

8 Research article

9 Running title: Recombination-stable multimeric GFP

10

11 Authors:

12 Peter Rugbjerg¹ – petru@biosustain.dtu.dk

13 Christoph Knuf¹ – chrik@biosustain.dtu.dk

14 Jochen Förster¹ – jfor@biosustain.dtu.dk

15 Morten O. A. Sommer¹* - msom@bio.dtu.dk

16

17 1) Novo Nordisk Foundation Center for Biosustainability, Technical University of

18 Denmark

19 Kogle Allé 6, DK-2970 Hørsholm, Denmark

20

21 * Corresponding author

22

23 **Abstract**

24 Green fluorescent proteins (GFPs) are widely used for visualization of proteins to
25 track localization and expression dynamics. However, phenotypically important
26 processes can operate at too low expression levels for routine detection, i.e. be
27 overshadowed by autofluorescence noise. While GFP functions well in translational
28 fusions, the use of tandem GFPs to amplify fluorescence signals is currently avoided
29 in *Saccharomyces cerevisiae* and many other microorganisms due to the risk of loop-
30 out by direct-repeat recombination. We increased GFP fluorescence by translationally
31 fusing three different GFP variants, yeast-enhanced GFP, GFP+ and superfolder GFP
32 to yield a sequence-diverged triple GFP molecule 3vGFP with 74-84 % internal repeat
33 identity. Unlike a single GFP, the brightness of 3vGFP allowed characterization of a
34 weak promoter in *S. cerevisiae*. Utilizing 3vGFP, we further engineered a less leaky
35 Cu²⁺-inducible promoter based on *CUPI*. The basal expression level of the new
36 promoter was approx. 61 % below the wild-type *CUPI* promoter, thus expanding the
37 absolute range of Cu²⁺-based gene control. The stability of 3vGFP towards direct-
38 repeat recombination was assayed in *S. cerevisiae* cultured for 25 generations under
39 strong and slightly toxic expression after which only limited reduction in fluorescence
40 was detectable. Such non-recombinogenic GFPs can help quantify intracellular
41 responses operating a low copy number in recombination-prone organisms.

42

43 **Keywords:** signal amplification, synthetic biology, promoter engineering, protein
44 multimerization

45

46 **Introduction**

47 Green fluorescent protein (GFP) is an invaluable tool for real-time visualization of
48 intracellular proteins. Since the initial cloning, numerous improvements, variants and
49 applications have been developed (Snapp 2009; Miyawaki 2011). GFP is particularly
50 useful for quantification of intracellular events, localizations and populations at
51 single-cell resolution. However, a minimal expression level is required such that the
52 fluorescent output exceeds the cell autofluorescence and produces detectable signals.
53 Still, biologically important processes occur through the interaction of a few
54 molecules per cell, which is hard to quantify using existing fluorescent proteins and
55 non-specialized experimental setups (Raj and van Oudenaarden 2009; Li and Xie
56 2011; Gahlmann and Moerner 2014). Further, the engineering of synthetic cell
57 functionalities can depend on fine characterization and balancing of low gene
58 expression levels (Ajikumar et al. 2010; Harton et al. 2013).

59 The strategies for improving fluorescent output signals include the design of new GFP
60 variants such as GFP⁺, yeast-enhanced GFP (yEGFP) and superfolder GFP (sfGFP)
61 (Cormack et al. 1997; Scholz et al. 2000; Pédelacq et al. 2006). Still, monitoring of
62 single-molecule events such as chromosome movements in *Escherichia coli* has e.g.
63 required multimerization of 96 DNA binding sites to localize enough fluorescent
64 protein to produce a distinguishable signal (Xie et al. 2008). Artificial tethering of a
65 bright yellow fluorescent protein (Venus YFP) to the inside *E. coli* cell membrane
66 allowed a microscope-detectable signal from a single YFP-tagged protein (Yu et al.
67 2006). Thus without techniques for single-molecule GFP sensitivity, the full-genome
68 mapping of subcellular protein localization in *Saccharomyces cerevisiae* (yeastGFP)
69 did not produce signals above background for 361 open reading frames (8 pct. of

70 total) otherwise shown to be expressed in the growth phase assayed (Ghaemmaghami
71 et al. 2003; Huh et al. 2003). Equivalently, the issue of not detecting all low-
72 expressing *S. cerevisiae* proteins was also observed when the GFP library was applied
73 to flow cytometry (Newman et al. 2006).

74 In some contexts, simple overexpression may shed light over the lacking information,
75 but since the location of many proteins is a result of interactions with other cell
76 components, a radical change in copy number could easily result in artificial
77 observations. In other situations, the target output is the activity of specific weak
78 promoters, e.g. in synthetic biological circuits, fluorescence-coupled biosensors or
79 when developing promoter libraries. Several technologies permit the engineering of
80 new promoters, e.g. responsive to other inducer molecules by hybridizing with
81 upstream TF-binding sites (Blazek and Alper 2013) or tuned to match fine, desirable
82 transcription levels through mutagenesis of a strong native promoter (Nevoigt et al.
83 2006). Difficulties in GFP detection may have been a limitation in these
84 developments for weaker promoter levels, though low expression may be
85 phenotypically important for a wide range of synthetic biology purposes. In synthetic
86 circuit designs, any concealed information on the shape of dose-response curves
87 inhibits the analysis of mechanistic clues otherwise given by the response curvature
88 (Ang et al. 2013). In applications of metabolite biosensors, background-covered
89 signal levels means that the full regulatory capability cannot be utilized, e.g. limiting
90 subsequent fluorescence-activated cell sorting (FACS). Ultimately, such
91 autofluorescence could conceal properly functional GFP completely (Billinton and
92 Knight 2001).

93

94 The efforts aimed at reducing the autofluorescence target two phenomena: Simple
95 medium autofluorescence arises from measuring fluorescence without isolating cells
96 from medium, e.g. in continuously growing cultures. These effects can be reduced by
97 the choice of medium or spectral unmixing by correcting for autofluorescence from a
98 wavelength representing effects of the culture medium (Lichten et al. 2014).
99 However, the cell autofluorescence is a more central issue, i.a. resulting from the
100 fluorescence of flavins and NAD(P)H (Billinton and Knight 2001). Cellular
101 autofluorescence also impacts techniques such as flow cytometry and microscopy and
102 the weak signal intensity must be amplified intrinsically to the cell.
103
104 Previous studies in mammalian cell lines have tackled the obstacle of cell
105 autofluorescence using directly repeated GFPs typically fused three to six times in
106 tandem using a small translational linker (Genové et al. 2005). By such approaches, it
107 has been possible to achieve good linear increments in fluorescence signals. However,
108 tandem repeats are problematic in organisms with proficient homologous
109 recombination such as *Escherichia coli* or *S. cerevisiae* where recombination between
110 DNA can happen within windows of identity at around 25 nucleotides (Ahn et al.
111 1988). This could explain why tandem GFP methods are avoided in these organisms.
112 However, even slight sequence divergence between repeats substantially decreases the
113 rate of recombination as seen in the case of recombination between 350 bp inverted
114 repeats, which was 4,600-fold reduced when sequence identity was reduced from 100
115 % to 74 % in *S. cerevisiae* (Datta et al. 1997). Similar effects occur in *E. coli* where
116 up to 1,000-fold reduction was observed following a reduction in repeat identity to 80
117 % (Rayssiguier et al. 1989).

118 Thus, in this study we present a simple methodology to take advantage of the ability
119 to add sequence divergence to tandem proteins while maintaining function through
120 variation in amino acid sequence as well as synonymous codon usage. By fusing three
121 different GFP variants that vary mainly at nucleotide-level, we produce a new triple
122 tandem GFP (3vGFP) stabilized towards direct-repeat recombination. We
123 demonstrate the utility of 3vGFP through a genetically triggered promoter (ON/OFF)
124 and developing and characterizing a new version of a Cu²⁺-responsive promoter with
125 reduced leakiness. Application of 3vGFP allowed visualization of weak signals that
126 could not be separated from autofluorescence levels using the brightest individual
127 GFP variant, superfolder GFP. Lastly, we test the stability towards recombination
128 after culturing of the strain harboring 3vGFP through 25 generations.

129 **Materials and methods**

130 **Materials**

131 Unless otherwise stated, reagents were purchased from Sigma-Aldrich. Synthetic
132 complete (SC) medium was prepared from 1.4 g/L synthetic complete drop-out mix
133 lacking uracil, tryptophan, leucine and histidine (Y2001), 6.7 g/L yeast nitrogen base
134 without amino acids (Y0626) and 20 g/L D-glucose, pH standardized to 5.6. When SC
135 was supplemented with additional amino acids, 60 mg/L leucine, 20 mg/L uracil,
136 20 mg/L histidine-HCl and 20 mg/L tryptophan was added. Yeast Peptone Dextrose
137 medium contained 20 g/L D-glucose.

138 Oligonucleotides were purchased from Integrated DNA Technologies.

139 **Plasmids**

140 The plasmids employed in this study are listed in Table 1.

141 **Table 1** Plasmids employed in this study, describing whether they lead to
 142 chromosomal integration or propagate autonomously in *S. cerevisiae*.

Plasmid	Expression cassette (promoter-ORF-terminator)	Maintenance in <i>S. cerevisiae</i> through	Reference
pPR4-3vGFP	pSPAL10-3vGFP-tURA3	<i>CEN/ARS, HIS3</i>	This study
pPR4-sfGFP	pSPAL10-sfGFP-tURA3	<i>CEN/ARS, HIS3</i>	This study
pCU2-3vGFP	pCUP1dim -3vGFP- tURA3	<i>CEN/ARS, URA3</i>	This study
pCfB258-CUP1-3vGFP	pCUP1-3vGFP-tCYC1	Chromosomal integration	This study
pCfB258-CUP1-SPO13-3vGFP	pCUP1dim -3vGFP- tCYC1	Chromosomal integration	This study
pDS1U-X2-3vGFP	pTEF1-3vGFP	Chromosomal integration	This study
pEXP22	pADH1-GAL4AD-RalGDS-tADH1	<i>TRP1</i>	Life Technologies
pEXP32	pADH1-GAL4DBD-Krev1-tADH1	<i>LEU2</i>	Life Technologies
pRS413	-	<i>LEU2</i>	(Sikorski and Hieter, 1989)
pRS415	-	<i>HIS3</i>	(Sikorski and Hieter, 1989)

143

144 **Strains**

145 The strains analyzed in this study are listed in Table 2.

146 The following background strains were used to construct the strains:

147 *Saccharomyces cerevisiae* MaV203 (MAT α , *leu2-3,112*, *trp1-901*, *his3 Δ 200*, *ade2-*
 148 *101*, *gal4 Δ* , *gal80 Δ* , *SPAL10::URA3*, *GAL1::lacZ*, *HIS3UAS GAL1::HIS3@LYS2*,

149 *can1^R, cyh2^R* (Purchased from Life Technologies).

150 *Saccharomyces cerevisiae* PRa18 (*MAT α , leu2-3,112, trp1-901, his3 Δ 200, ade2-101,*
151 *gal4 Δ , gal80 Δ , GAL1::lacZ, can1^R, cyh2^R*) Derived from *S. cerevisiae* MaV203.

152 *Saccharomyces cerevisiae* PRa26: *MAT α , leu2-3,112, trp1-901, his3 Δ 200, ade2-101,*
153 *gal4 Δ , gal80 Δ , GAL1::lacZ, rad16::KanMX, can1^R, cyh2^R*. Derived from *S. cerevisiae*
154 PRa18.

155 *Saccharomyces cerevisiae* CfB1010 (*MAT α ; ura3-52; his3 Δ 1; leu2-3/112; MAL2-8^c;*
156 *SUC2; are2 Δ ::loxP-KanMX; X-3::tHMG1-P_{TEF1}-P_{PGK1}-AtATR2*). Derived from *S.*
157 *cerevisiae* CEN.PK 102-5B.

158

159 **Table 2** *S. cerevisiae* strains analyzed in this study, indicating which plasmids or
 160 chromosomal integrations were introduced into the respective parental strains.

Strain name	Promoter	GFP	Plasmid #1	Plasmid #2	Plasmid #3	Integrative plasmid	Parent strain
PRa106	ON	3vGFP	pPR4-3vGFP	pEXP32	pEXP22	-	PRa26
PRa107	OFF	3vGFP	pPR4-3vGFP	pRS415	pEXP22	-	PRa26
PRa108	-	-	pRS413	pRS415	pEXP22	-	PRa26
PRa109	ON	sfGFP	pPR4-sfGFP	pEXP32	pEXP22	-	PRa26
PRa110	OFF	sfGFP	pPR4-sfGFP	pRS415	pEXP22	-	PRa26
CK24	pCUP1	3vGFP	-	-	-	pCfB258-CUP1-3vGFP	CfB1010
CK28	pCUP1dim	3vGFP	-	-	-	pCfB258-CUP1-SPO13-3vGFP	CfB1010
PRa114	pTEF1	3vGFP	-	-	-	pDS1U-X2-3vGFP	PRa18

161

162 **Construction of 3vGFP plasmids**

163 Plasmids were constructed by uracil-excision (USER) cloning. The general method
 164 for USER cloning was based on agarose gel-purification of the PCR products
 165 amplified using DNA polymerase X7 (Nørholm 2010). These were mixed in an
 166 equimolar 20 μ L reaction with 0.5 μ L USER enzyme (New England Biolabs) and 0.5

167 μ L DpnI FastDigest (Thermo Scientific) in FastDigest buffer at 37 degrees C for 1-2
168 hours. Following 25 minutes at room temperature, 2.5 μ L reaction was transformed
169 into *E. coli*. Correctly cloned plasmids were identified using restriction analysis and
170 DNA sequencing. The detailed use of oligonucleotides for assembly of all plasmids is
171 described in Supplementary data.

172 **Construction of strains**

173 Plasmids and DNA for chromosomal targeting was introduced in *S. cerevisiae* by
174 methods described previously (Gietz and Schiestl 2007). The PRa18 strain was
175 constructed from the MaV203 strain by deletion of *SPAL10::URA3* through
176 replacement with a *kanMX* gene deletion cassette flanked by loxP recombination sites
177 from the pUG6 plasmid as described before (Güldener et al. 1996). DNA flanks to
178 direct homologous recombination of the cassette to the chromosomal locus were
179 generated by PCR on *S. cerevisiae* MaV203 gDNA spanning a fragment from 5'-
180 CCATTCAACTAACATCACAC to 5'-CCTTCACCATAAATATGCC (upstream
181 flank) and from 5'-CTCACAAATTAGAGCTTC to 5'-CCCATATCCAACCTCCAA
182 (downstream flank). These flanks were cloned to the *kanMX* gene deletion cassette
183 and transformed into yeast. The *kanMX* cassette was looped out by heterologous
184 expression of Cre recombinase from the pSH47 plasmid (Güldener et al. 1996). To
185 construct PRa26 subsequently, the chromosomal *HIS3* gene within the *rad16* locus
186 was deleted using the same *kanMX* approach. The targeting flanks spanned regions
187 from 5'- AGTTGGTACACCAGTTATACGG to 5'-
188 AAAGCATAGGATACCGAGAAAC (upstream flank) and 5'-
189 TGACATCACCCGAAAAGAAGC to 5'- GATTATGGTTACGATGTCGA
190 (downstream flank).

191 To construct PRa114, the pTEF1-3vGFP construct was chromosomally integrated into
192 the PRa18 strain using divisible selection (Rugbjerg et al. 2015). DNA fragments for
193 integration was liberated from the vector pDS1U-X2-3vGFP by digestion with *SmiI*
194 and transformed into yeast along with empty divisible selection plasmids pDS2 and
195 pDS3 in order to reconstitute the selectable Ura⁺ phenotype.

196 To construct respectively CK24 and CK28 from the Cfb1010 strain, the pCUP1-
197 3vGFP and pCUP1dim-3vGFP was chromosomally integrated by cloning into the
198 EasyClone integrative vectors (Jensen et al. 2013). The DNA fragments for
199 integration were obtained through *NotI* digestion of the vectors pCfB258-CUP1-
200 3vGFP and pCfB258-CUP1-SPO13-3vGFP respectively, followed by agarose gel
201 purification.

202 **Estimation of TEF1-3vGFP fitness cost**

203 Microtiter cultures of 200 μ L YPD was inoculated by 100x backdilution of overnight
204 YPD pre-cultures of PRa114 and PRa108, each inoculated from single colonies. The
205 cultures were cultivated in a 96-well plate at 30 deg. C and continuous shaking in an
206 ELx808 plate reader (BioTek), set to measure optical density every 15 minutes at
207 OD₆₃₀. The plate was covered with a BreathSeal (Greiner Bio-one) and plastic lid.
208 Growth rates were calculated for all three biological replicates by exponential
209 regression between OD₆₃₀ and time (hours) during the same OD₆₃₀ span of
210 exponential growth phase. All OD₆₃₀ values were initially standardized to the time
211 zero reading to account for differences in seal absorbance.

212 **Cultivations for stability tests**

213 The PRa114 strain was cultured from a single colony inoculated in 25 mL YPD
214 medium and cultured at 30 deg. C and 250 rpm horizontal shaking in three parallel

215 lineages. By measuring OD₆₀₀, the number of generations passed was calculated. 2 %
216 of the culture was passed to fresh medium and grown again until totally 25
217 generations had passed. For comparison between cultured population and reference
218 strain, approx. 25 µL of each cell population was inoculated in YPD medium at the
219 same time and cultured at 30 deg. C for 16 hours with 250 rpm horizontal shaking.

220 **Fluorescence measurements**

221 Pre-cultures in selective SC medium were inoculated from single colonies and
222 cultures overnight at 30 deg. C. From these, 200 µL microtiter cultures of selective
223 SC medium were inoculated and cultured at 30 deg. C with 300 rpm horizontal
224 shaking in an Innova shaking incubator for 16 hours. As cover, the microtiter plates
225 were covered with a BreathSeal (Greiner Bio-one) and a plastic lid.
226 The cell cultures were diluted approx. 1:100 in FACS flow buffer (BD Biosciences)
227 and analyzed on a LSR Fortessa flow cytometer (BD Biosciences) equipped with a
228 blue laser (488 nm) and set to measure 10,000 cells within a gate defined by forward
229 and side scatter to capture all yeast cells. A FITC filter (530/30 nm) was used to
230 measure GFP fluorescence reporting the area of the measured peaks. The laser voltage
231 was adjusted to optimally utilize the dynamic range of detection. Data was processed
232 and visualized as histograms with FlowJo version 10 (default settings) by overlaying
233 the populations for each particular comparison.

234

235 **Sequence alignment**

236 Simple nucleotide and protein sequence alignment was performed using the ClustalO
237 algorithm (Sievers et al. 2011).

238

239 **Results and discussion**

240 **Amplification of fluorescence by tandems of differently encoded GFPs**

241 To amplify the fluorescence signal of a GFP molecule while keeping transcription
242 strength constant, the new 3vGFP protein was engineered by fusion of nucleotide
243 sequences encoding yEGFP, GFP+ and sfGFP (Cormack et al. 1997; Pédelacq et al.
244 2006) (Fig. 1A). Two glycine residues were introduced as translational linker in each
245 junction. The fluorescence of 3vGFP was evaluated when expressed from a weak *S.*
246 *cerevisiae* hybrid promoter (p*SPALI0*) (Vidal et al. 1996) based on p*SPO13* to mimic
247 low-expression applications (Huang and Schreiber 1997; Harton et al. 2013). The
248 low-level strength of p*SPALI0* is attained by utilizing the UME6 repressor binding
249 site naturally present within the *SPO13* promoter, which allows very low expression
250 levels e.g. useful for control of cell growth. Further, GAL4-binding sites fused 179 bp
251 upstream of start codon provide an upstream activating sequence, allowing
252 transcription factor-based ON/OFF inputs.

253 The output fluorescence was first evaluated with single sfGFP (Fig. 1B), which is the
254 individually brightest of the three GFPs tested. However, the fluorescence levels
255 could not be distinguished from the control strain devoid of genes encoding GFP
256 (PRa108). In contrast, the fluorescence of a strain (PRa106) carrying the gene
257 encoding 3vGFP controlled by the same promoter was 3-fold higher than the
258 background level and thus the level of the single sfGFP strain (Fig. 1B).

259

260 To test the utility of 3vGFP as output signal in a synthetic biology setting, we
261 constructed versions of the strain with the p*SPALI0* promoter turned OFF. The

262 promoter is activated (ON) when a hybrid GAL4 activation domain binds a cognate
263 hybrid GAL4 DNA-binding domain, which interacts with GAL4-binding sites of
264 p*SPAL10*. The protein-protein interaction domains were based on the known Krev1
265 and RalGDS interaction domains (Herrmann et al. 1996). However omitting the
266 DNA-binding domain prevents reconstitution of a functional transactivator (OFF).
267 These ON/OFF effects of present DNA-binding domain remained hidden below the
268 background levels of the sfGFP strains, while observable in strains with 3vGFP as
269 output (Fig. 1B).

270

271 **Figure 1**

272 **Stability towards recombination**

273 Direct-repeat recombination in mitotic *S. cerevisiae* is reported to occur at rates
274 between $5.8 \cdot 10^{-5}$ and $12 \cdot 10^{-5}$ per cell generation for repeats of several kilo base pair
275 identity (Dornfeld and Livingston 1992). This recombination rate is linearly
276 dependent on identity length at such long segments, however the rate drops rapidly
277 below the minimal efficient processing segment (MEPS) length at around 250 bp in *S.*
278 *cerevisiae* (Jinks-Robertson et al. 1993). While internal identity of 3vGFP ranges 74-
279 84 % (Fig. 2B), the identical segments are maximally at a ten-fold shorter length than
280 the MEPS.

281 To test the recombination stability of 3vGFP, we wanted to measure whether the
282 fluorescence levels originating from 3vGFP would attenuate following repeated
283 culturing. While the 3vGFP molecule is engineered to limit direct-repeat
284 recombination, long-term cultivation could potentially still lead to this especially if
285 favored by a concurrent fitness advantage. To test stability at high expression level,
286 we therefore also chromosomally integrated *3vGFP* under control of the strong

287 promoter from *TEF1* i.e. at a level surpassing the intended use of 3vGFP. Expressing
288 3vGFP from the *TEF1* promoter caused a considerable cost in fitness of approx. 15 %
289 in YPD, reducing the growth rate from an average of 0.35 hr⁻¹ to 0.30 hr⁻¹ compared
290 to the negative control strain PRa108. Following culturing by serial passing (2 %) of
291 liquid cultures for 25 generations of three parallel lineages, single-cell level analysis
292 revealed that the average fluorescence level of the cell population had diminished by
293 7 percent, perhaps due to spontaneous direct-repeat recombination. The single cell-
294 level visualization indicated a slight left-shift of the population (Fig 2A). These
295 results exemplify that direct-repeat recombination can occur within 3vGFP in *S.*
296 *cerevisiae* and if selected for, these effects can become significant. However, since
297 3vGFP is intended for use at levels of low expression, a fitness advantage is not likely
298 to further drive diminished fluorescence at a typical utility of 3vGFP.

299

300 **Figure 2.**

301

302 **Application of 3vGFP to construct an inducible promoter with reduced leakiness**

303 Inducible promoters are important for development of e.g. synthetic genetic circuits,
304 but the leakiness levels can be problematic in certain uses. To demonstrate the utility
305 of 3vGFP, we therefore wanted to use it as output for genetic re-engineering of the
306 popular Cu^{2+} -responsive promoter of *S. cerevisiae* *CUPI*. *pCUI* has been employed
307 in many different biotechnological cases (Labbé and Thiele 1999; Scholz et al. 2000;
308 Rugbjerg et al. 2013), but displays considerable baseline activity (leakiness). *pCUI*
309 induction results from elevated Cu^{2+} concentrations mediated through binding of Cu^{2+}
310 to the ACE1 transcription factor, which in turn binds to upstream activating sequence
311 (UAS) elements of *pCUI* (Huibregtse 1989; Evans et al. 1990) (elements
312 schematically depicted in Fig. 3A). The leakiness level of *pCUI* measured with
313 3vGFP corresponded to 2.5-fold the cell autofluorescence (Fig. 3B). Based on the
314 regulatory mechanism of ACE1, we anticipated that trace levels of Cu^{2+} in the growth
315 medium did not cause this leakiness, but rather assumed this basal transcriptional
316 activity to be ACE1-independent. Accordingly, as strategy we hypothesized that
317 swapping the promoter region downstream of ACE1 UASs for a transcriptionally
318 repressed promoter could provide attenuation, while maintaining the response to
319 ACE1-dependent induction. We therefore combined the upstream region of *pCUI* (-
320 149 to -454) containing three ACE1-binding sites, with part of the *S. cerevisiae*
321 *pSPO13* (-1 to -157) including its UME6 repressor-binding site (Fig. 3A). This new
322 promoter (*pCUI_{dim}*) controlling 3vGFP resulted in fluorescence that was reduced
323 approx. 61 % (before background-subtraction) to levels close to the cell
324 autofluorescence (Fig. 3B), while the promoter remained responsive to addition of
325 Cu^{2+} (Fig. 3C).

326

327 **Figure 3**

328

329 The recombination-stabilized tandem GFP described in this study can enable
330 characterization of minimally expressed genes in recombination-efficient organisms
331 such as *S. cerevisiae* and other yeasts. As shown in this study, 3vGFP allowed
332 characterization of the activation of a weak promoter and accordingly characterization
333 of manipulations taking place at such low expression levels. Further, this particular
334 approach of recombination-stabilizing GFPs with different protein and nucleotide
335 sequences can be scaled in number. Recent brighter fluorescent proteins could be
336 applied such as mNeonGreen (Shaner et al. 2013).

337 In principle, sequence divergence could be generated strictly at nucleotide level
338 through codon optimization of segments encoding the same protein. Codon
339 optimization can however introduce significant effects on the translation efficiencies
340 (Goodman et al. 2013). Another concern may be spurious promoter/RBS activities,
341 which could theoretically cause transcription and translation initiation from locations
342 within the tandem GFP, thus producing truncated tandem proteins. Such situations
343 would complicate the isolation of promoter responses and might require alleviation of
344 the second and third GFP start codon.

345 An alternative method for assessment of promoter activities could be the use of the
346 fluorescent RNA of the Spinach family, which bypasses the step of translation since
347 the RNA forms the fluorescent signal (Paige et al. 2012; Pothoulakis et al. 2014).

348 However, while the technology has potential for synthetic biological use, its general
349 applicability remains to be seen, such as the detection limits for low expression levels.

350 Further relevant, fluorescent *in situ* hybridization for RNA (RNA FISH) is a

351 technique allowing sensitive detection of transcripts at single-cell level (Zenklusen et
352 al. 2008). This alleviates genetic engineering, but entails more sample treatment than
353 for detection of GFP fluorescence.

354 In this study, a new simple strategy for engineering tandem fluorescent proteins was
355 employed to produce brighter GFP signals with improved stability towards loop-out
356 recombination. GFPs with sequence variation mainly at nucleotide level were
357 translationally linked to form a recombination-stabilized tandem GFP molecule
358 3vGFP. Such GFPs could be useful for characterizing promoter activities in the range
359 where normal single GFP signals fall below the cell autofluorescence levels. We
360 specifically applied the 3vGFP molecule to characterize the ON/OFF levels of a weak
361 promoter, which was not possible using a single sfGFP, and to develop a new hybrid
362 Cu^{2+} -responsive promoter pCUP1dim with lower leakiness level. The plasmid pCU2-
363 3vGFP encompassing the nucleotide sequence of 3vGFP and pCUP1dim will be
364 deposited at the Addgene repository.

365 **Competing interests**

366 The authors declare that they have no competing interests.

367 **Funding**

368 This work was supported by the Novo Nordisk Foundation, the European Union
369 Seventh Framework Programme (FP7-KBBE-2013-7-single-stage) under Grant
370 agreement no. 613745, Promys, and Deutsche Bundesstiftung Umwelt,.

371 **Acknowledgement**

372 George Church is acknowledged for sfGFP encoded on pJ251-GERC (AddGene
373 plasmid 47441).

374

375

376

377

378 References

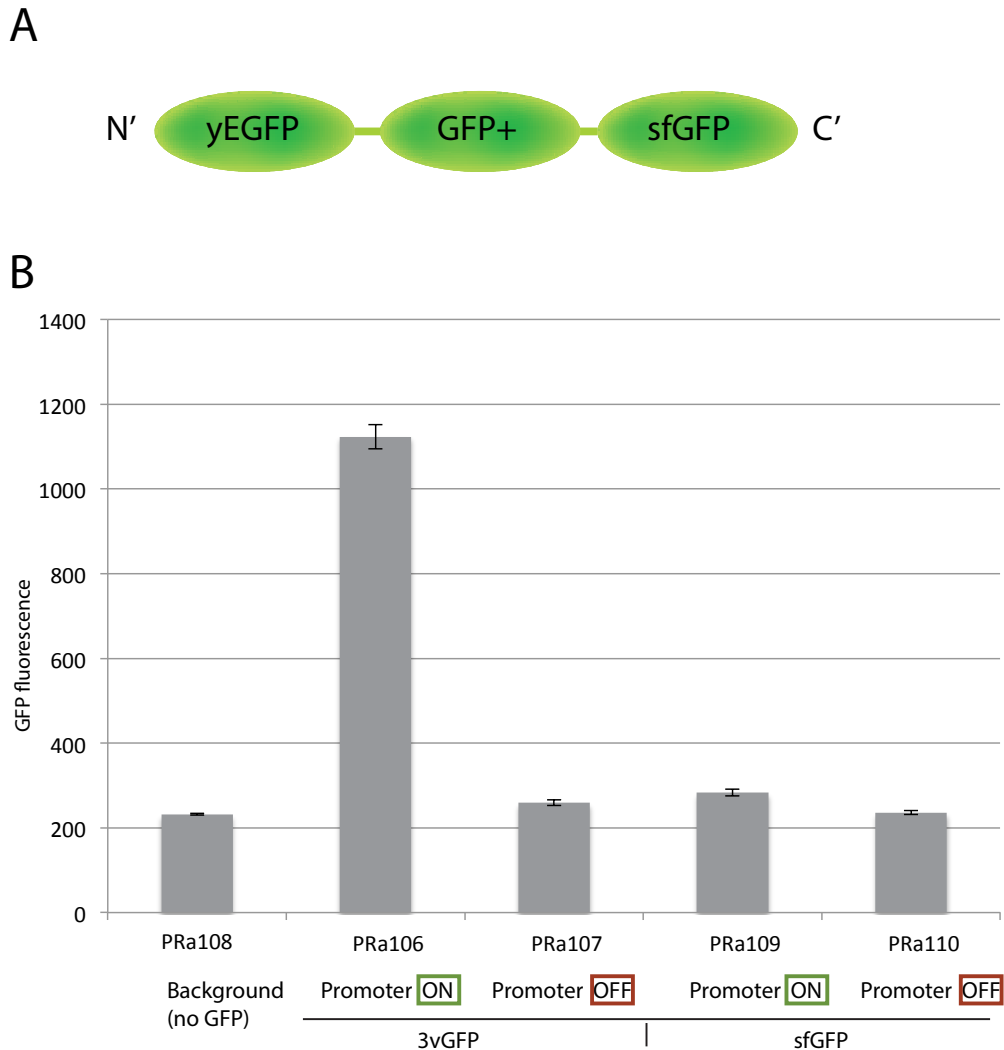
- 379 Ahn BY, Dornfeld KJ, Fagrelus TJ, Livingston DM. 1988. Effect of limited homology on gene
380 conversion in a *Saccharomyces cerevisiae* plasmid recombination system. *Mol. Cell. Biol.*
381 8:2442–2448.
- 382 Ajikumar PK, Xiao W-H, Tyo KEJ, Wang Y, Simeon F, Leonard E, Mucha O, Phon TH, Pfeifer B,
383 Stephanopoulos G. 2010. Isoprenoid pathway optimization for Taxol precursor overproduction in
384 *Escherichia coli*. *Science* 330:70–74.
- 385 Ang J, Harris E, Hussey B. 2013. Tuning Response Curves for Synthetic Biology. *ACS Synth. Biol.*
386 [Internet]:547–567. Available from: <http://pubs.acs.org/doi/abs/10.1021/sb4000564>
- 387 Billinton N, Knight a W. 2001. Seeing the wood through the trees: a review of techniques for
388 distinguishing green fluorescent protein from endogenous autofluorescence. *Anal. Biochem.*
389 [Internet] 291:175–197. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/11401292>
- 390 Blazeck J, Alper HS. 2013. Promoter engineering: Recent advances in controlling transcription at the
391 most fundamental level. *Biotechnol. J.* 8:46–58.
- 392 Cormack B, Bertram G, Egerton M. 1997. Yeast-enhanced green fluorescent protein (yEGFP): a
393 reporter of gene expression in *Candida albicans*. *Microbiology* [Internet] 143:303–311. Available
394 from: <http://mic.sgmjournals.org/content/143/2/303.short>
- 395 Datta A, Hendrix M, Lipsitch M, Jinks-Robertson S. 1997. Dual roles for DNA sequence identity and
396 the mismatch repair system in the regulation of mitotic crossing-over in yeast. *Proc. Natl. Acad.*
397 *Sci.* [Internet] 94:9757–9762. Available from: <http://www.pnas.org/content/94/18/9757.short>
- 398 Dornfeld KJ, Livingston DM. 1992. Plasmid recombination in a *rad52* mutant of *Saccharomyces*
399 *cerevisiae*. *Genetics* 131:261–276.
- 400 Evans C, Engelke D, Thiele D. 1990. ACE1 Transcription Factor Produced in *Escherichia coli* Binds
401 Multiple Regions within Yeast Metallothionein Upstream Activation site Sequences. *Mol. Cell.*
402 *Biol.* [Internet] 10:426–429. Available from: <http://mcb.asm.org/content/10/1/426.short>
- 403 Gahlmann A, Moerner WE. 2014. Exploring bacterial cell biology with single-molecule tracking and
404 super-resolution imaging. *Nat. Rev. Microbiol.* [Internet] 12:9–22. Available from:
405 <http://www.ncbi.nlm.nih.gov/pubmed/24336182>
- 406 Genové G, Glick B, Barth A. 2005. Brighter reporter genes from multimerized fluorescent proteins.
407 *Biotechniques* [Internet] 39:814–822. Available from:
408 <http://www.biotechniques.com/article/05396BM02>
- 409 Ghaemmaghami S, Huh W-K, Bower K, Howson RW, Belle A, Dephoure N, O’Shea EK, Weissman
410 JS. 2003. Global analysis of protein expression in yeast. *Nature* [Internet] 425:737–741.
411 Available from: <http://www.ncbi.nlm.nih.gov/pubmed/14562106>
- 412 Gietz RD, Schiestl RH. 2007. High-efficiency yeast transformation using the LiAc/SS carrier
413 DNA/PEG method. *Nat. Protoc.* [Internet] 2:31–34. Available from:
414 <http://www.ncbi.nlm.nih.gov/pubmed/17401334>
- 415 Goodman D, Church G, Kosuri S. 2013. Causes and effects of N-Terminal Codon Bias in Bacterial
416 Genes. *Science* (80-.). [Internet] 342:475–479. Available from:
417 <http://www.sciencemag.org/content/342/6157/475.short>

- 418 Güldener U, Heck S, Fielder T, Beinhauer J, Hegemann JH. 1996. A new efficient gene disruption
419 cassette for repeated use in budding yeast. *Nucleic Acids Res.* [Internet] 24:2519–2524.
420 Available from:
421 [http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=145975&tool=pmcentrez&rendertype](http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=145975&tool=pmcentrez&rendertype=abstract)
422 [=abstract](http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=145975&tool=pmcentrez&rendertype=abstract)
- 423 Harton M, Wingler L, Cornish V. 2013. Transcriptional Regulation Improves the Throughput of Three-
424 Hybrid Counter Selections in *Saccharomyces cerevisiae*. *Biotechnol. J.* [Internet]:1–23.
425 Available from: <http://onlinelibrary.wiley.com/doi/10.1002/biot.201300186/abstract>
- 426 Herrmann C, Horn G, Spaargaren M, Wittinghofer A. 1996. Differential Interaction of the Ras Family
427 GTP-binding Proteins H-Ras, Rap1A, and R-Ras with the Putative Effector Molecules Raf
428 Kinase and Ral-Guanine Nucleotide Exchange Factor. *J. Biol. Chem.* [Internet] 271:6794–6800.
429 Available from: <http://www.jbc.org/cgi/doi/10.1074/jbc.271.12.6794>
- 430 Huang J, Schreiber SL. 1997. A yeast genetic system for selecting small molecule inhibitors of protein-
431 protein interactions in nanodroplets. *Proc. Natl. Acad. Sci. U. S. A.* [Internet] 94:13396–13401.
432 Available from:
433 [http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=28315&tool=pmcentrez&rendertype](http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=28315&tool=pmcentrez&rendertype=abstract)
434 [=abstract](http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=28315&tool=pmcentrez&rendertype=abstract)
- 435 Huh W-K, Falvo J V, Gerke LC, Carroll AS, Howson RW, Weissman JS, O’Shea EK. 2003. Global
436 analysis of protein localization in budding yeast. *Nature* [Internet] 425:686–691. Available from:
437 <http://www.ncbi.nlm.nih.gov/pubmed/14562095>
- 438 Huibregtse J. 1989. Copper-induced binding of cellular factors to yeast metallothionein upstream
439 activation sequences. *Proc. Natl. Acad. Sci.* [Internet] 86:65–69. Available from:
440 <http://www.pnas.org/content/86/1/65.short>
- 441 Jensen NB, Strucko T, Kildegaard KR, David F, Maury J, Mortensen UH, Forster J, Nielsen J,
442 Borodina I. 2013. EasyClone: method for iterative chromosomal integration of multiple genes in
443 *Saccharomyces cerevisiae*. *FEMS Yeast Res.* [Internet]:1–11. Available from:
444 <http://www.ncbi.nlm.nih.gov/pubmed/24151867>
- 445 Jinks-Robertson S, Michelitch M, Ramcharan S. 1993. Substrate length requirements for efficient
446 mitotic recombination in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* 13:3937–3950.
- 447 Labbé S, Thiele D. 1999. Copper ion inducible and repressible promoter systems in yeast. *Methods*
448 *Enzymol.* [Internet] 306:145–153. Available from:
449 <http://www.sciencedirect.com/science/article/pii/S0076687999060103>
- 450 Li G-W, Xie XS. 2011. Central dogma at the single-molecule level in living cells. *Nature* [Internet]
451 475:308–315. Available from:
452 [http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3600414&tool=pmcentrez&rendertype](http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3600414&tool=pmcentrez&rendertype=abstract)
453 [pe=abstract](http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3600414&tool=pmcentrez&rendertype=abstract)
- 454 Lichten C a, White R, Clark IBN, Swain PS. 2014. Unmixing of fluorescence spectra to resolve
455 quantitative time-series measurements of gene expression in plate readers. *BMC Biotechnol.*
456 [Internet] 14:11. Available from:
457 [http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3917901&tool=pmcentrez&rendertype](http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3917901&tool=pmcentrez&rendertype=abstract)
458 [pe=abstract](http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3917901&tool=pmcentrez&rendertype=abstract)
- 459 Miyawaki A. 2011. Proteins on the move: insights gained from fluorescent protein technologies. *Nat.*
460 *Rev. Mol. Cell Biol.* [Internet] 12:656–668. Available from:
461 <http://www.ncbi.nlm.nih.gov/pubmed/21941275>

- 462 Nevoigt E, Kohnke J, Fischer CR, Alper H, Stahl U, Stephanopoulos G. 2006. Engineering of promoter
463 replacement cassettes for fine-tuning of gene expression in *Saccharomyces cerevisiae*. *Appl.*
464 *Environ. Microbiol.* 72:5266–5273.
- 465 Newman JRS, Ghaemmaghami S, Ihmels J, Breslow DK, Noble M, DeRisi JL, Weissman JS. 2006.
466 Single-cell proteomic analysis of *S. cerevisiae* reveals the architecture of biological noise. *Nature*
467 441:840–846.
- 468 Nørholm MHH. 2010. A mutant Pfu DNA polymerase designed for advanced uracil-excision DNA
469 engineering. *BMC Biotechnol.* [Internet] 10:21. Available from:
470 <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2847956&tool=pmcentrez&rendertype=abstract>
471
- 472 Paige JS, Nguyen-Duc T, Song W, Jaffrey SR. 2012. Fluorescence imaging of cellular metabolites with
473 RNA. *Science* [Internet] 335:1194. Available from:
474 <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3303607&tool=pmcentrez&rendertype=abstract>
475
- 476 Pédelacq J-D, Cabantous S, Tran T, Terwilliger TC, Waldo GS. 2006. Engineering and characterization
477 of a superfolder green fluorescent protein. *Nat. Biotechnol.* [Internet] 24:79–88. Available from:
478 <http://www.ncbi.nlm.nih.gov/pubmed/16369541>
- 479 Pothoulakis G, Ceroni F, Reeve B, Ellis T. 2014. The spinach RNA aptamer as a characterization tool
480 for synthetic biology. *ACS Synth. Biol.* [Internet] 3:182–187. Available from:
481 <http://www.ncbi.nlm.nih.gov/pubmed/23991760>
- 482 Raj A, van Oudenaarden A. 2009. Single-Molecule Approaches to Stochastic Gene Expression. *Annu*
483 *Rev Biophys.* 255–270.
- 484 Rayssiguier C, Thaler DS, Radman M. 1989. The barrier to recombination between *Escherichia coli*
485 and *Salmonella typhimurium* is disrupted in mismatch-repair mutants. *Nature* 342:396–401.
- 486 Rugbjerg P, Myling-Petersen N, Sommer M. 2015. Flexible metabolic pathway construction using
487 modular and divisible selection gene regulators. *Metab. Eng.*:in press. Available from:
488 <http://www.sciencedirect.com/science/article/pii/S1096717615001019>
- 489 Rugbjerg P, Naesby M, Mortensen UH, Frandsen RJ. 2013. Reconstruction of the biosynthetic pathway
490 for the core fungal polyketide scaffold rubrofusarin in *Saccharomyces cerevisiae*. *Microb. Cell*
491 *Fact.* [Internet] 12:31. Available from:
492 <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3654996&tool=pmcentrez&rendertype=abstract>
493
- 494 Scholz O, Thiel A, Hillen W, Niederwieser M. 2000. Quantitative analysis of gene expression with an
495 improved green fluorescent protein. *Eur. J. Biochem.* [Internet] 267:1565–1570. Available from:
496 <http://onlinelibrary.wiley.com/doi/10.1046/j.1432-1327.2000.01170.x/full>
- 497 Shaner NC, Lambert GG, Chamma A, Ni Y, Cranfill PJ, Baird M a, Sell BR, Allen JR, Day RN,
498 Israelsson M, et al. 2013. A bright monomeric green fluorescent protein derived from
499 *Branchiostoma lanceolatum*. *Nat. Methods* [Internet] 10:407–409. Available from:
500 <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3811051&tool=pmcentrez&rendertype=abstract>
501
- 502 Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M,
503 Söding J, et al. 2011. Fast, scalable generation of high-quality protein multiple sequence
504 alignments using Clustal Omega. *Mol. Syst. Biol.* [Internet] 7:539. Available from:
505 <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3261699&tool=pmcentrez&rendertype=abstract>
506

- 507 Snapp EL. 2009. Fluorescent proteins: a cell biologist's user guide. *Trends Cell Biol.* [Internet]
508 19:649–655. Available from:
509 <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2784028&tool=pmcentrez&rendertype=abstract>
510
- 511 Vidal M, Brachmann RK, Fattaey a, Harlow E, Boeke JD. 1996. Reverse two-hybrid and one-hybrid
512 systems to detect dissociation of protein-protein and DNA-protein interactions. *Proc. Natl. Acad. Sci. U. S. A.* [Internet] 93:10315–10320. Available from:
513 <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=38381&tool=pmcentrez&rendertype=abstract>
514
515
- 516 Xie XS, Choi PJ, Li G-W, Lee NK, Lia G. 2008. Single-molecule approach to molecular biology in
517 living bacterial cells. *Annu. Rev. Biophys.* [Internet] 37:417–444. Available from:
518 <http://www.ncbi.nlm.nih.gov/pubmed/18573089>
- 519 Yu J, Xiao J, Ren X, Lao K, Xie X. 2006. Probing Gene Expression in Live Cells, One Protein
520 Molecule at a Time. *Science* (80-.). [Internet] 311:1600–1603. Available from:
521 <http://www.sciencemag.org/content/311/5767/1600.short>
- 522 Zenklusen D, Larson DR, Singer RH. 2008. Single-RNA counting reveals alternative modes of gene
523 expression in yeast. *Nat. Struct. Mol. Biol.* 15:1263–1271.
- 524
- 525

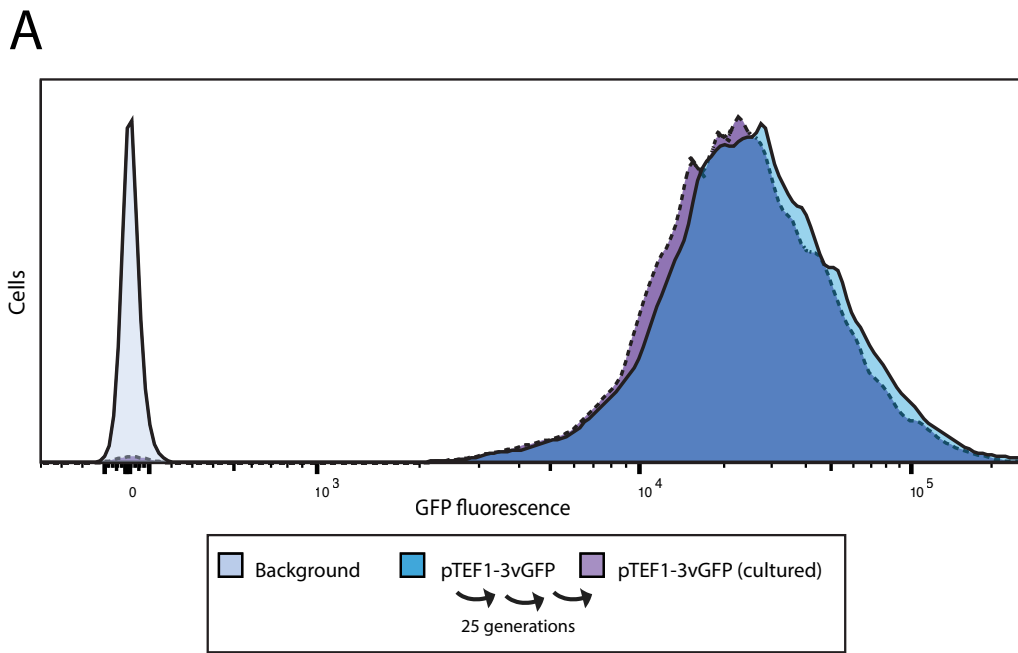
526 **Figures**



527

528 **Figure 1 Increased GFP fluorescence signal above autofluorescence level by**
 529 **triple tandem GFP (3vGFP).** A) Internal organization of individual GFP molecules
 530 fused as 3vGFP. 3vGFP consists of yeast-enhanced GFP (yEGFP), GFP+ and
 531 superfolder GFP. B) The *S. cerevisiae* strains carrying 3vGFP allowed the capture of
 532 the weak, ON/OFF promoter pSPAL10 unlike strains carrying a single sfGFP. The
 533 ON levels with single sfGFP corresponded to the background level of the empty
 534 control strain without GFP. The strains are described in detail in Table 2. Error bars
 535 depict standard error from biological replicates (n = 3).

536



B

Nucleotide-level identity

1: sfGFP	100.00		
2: yEGFP	74.23	100.00	
3: GFP+	76.33	84.45	100.00

Protein-level identity

1: sfGFP	100.00		
2: yEGFP	94.12	100.00	
3: GFP+	94.96	96.64	100.00

537

538 **Figure 2 Stability of the triple tandem GFP (3vGFP) towards loop-out**

539 **recombination.** A) Parallel lineages of a pTEF1-3vGFP *S. cerevisiae* strain was

540 cultured for 25 generations and re-measured to verify stability towards loop-out

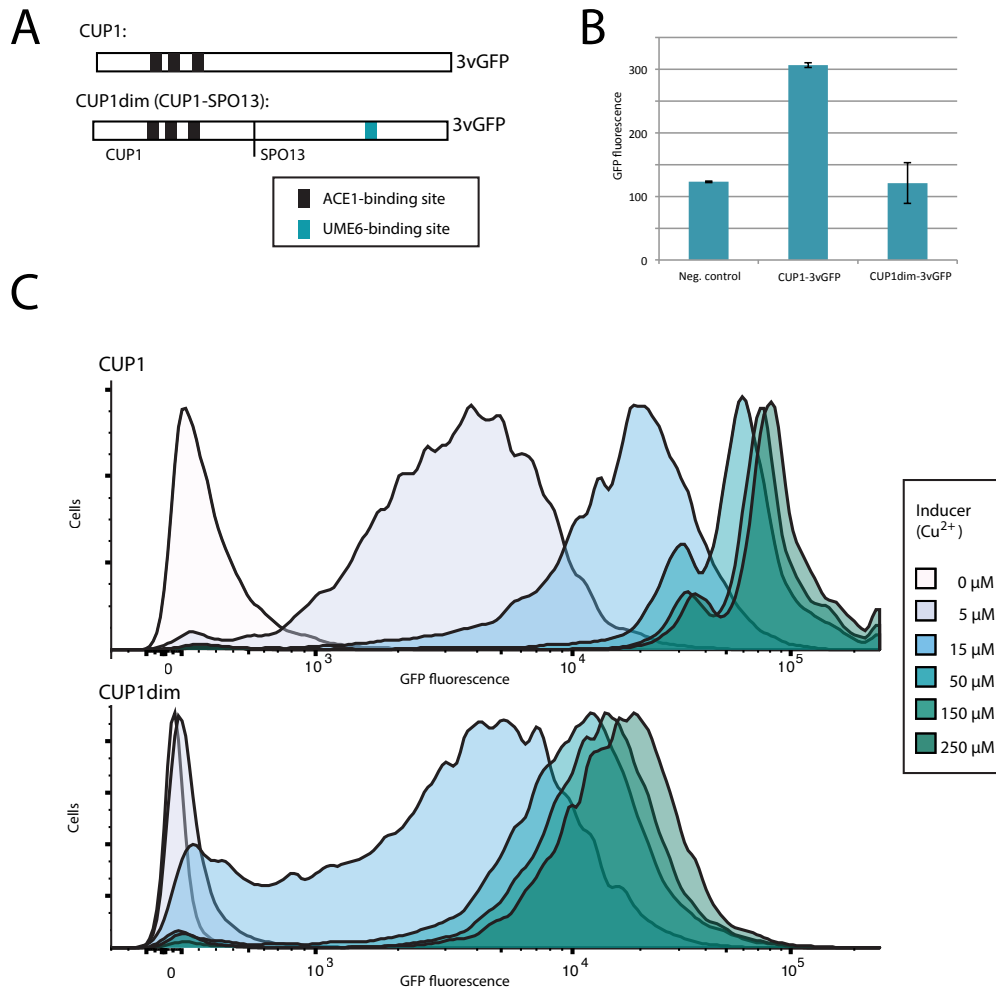
541 recombination, compared to a background strain without GFP. Flow cytometry of

542 representative example shown. Each sample contained 10,000 cells. The maxima of

543 the samples are standardized to an equal top point. B) Sequence identities between the

544 three direct repeats of sequences encoding GFP variants, as calculated by ClustalO.

545



546

547 **Figure 3 Development of weak Cu²⁺-responsive promoter through**

548 **characterization with 3vGFP.** A) Organization of DNA-binding sites for the Cu²⁺-

549 responsive ACE1 activator and UME6 repressor in the wildtype *CUP1* promoter and

550 the new dimmed, hybrid promoter p*CUP1dim*. B) OFF-level fluorescence measured

551 in absence of Cu²⁺ demonstrating the lower activity of the new hybrid promoter as

552 captured with 3vGFP. Error bars depict standard error from biological replicates (n =

553 3). C) Fluorescence of strain populations in response to addition of Cu²⁺. Flow

554 cytometry of representative example shown. Each sample contained 10,000 cells. The

555 maxima of the samples are standardized to an equal top point.

556