Title
Tunable Expression Systems for Orthogonal DNA Replication

Journal
ACS Synthetic Biology, 7(12)

ISSN
2161-5063

Authors
Zhong, Ziwei
Ravikumar, Arjun
Liu, Chang C

Publication Date
2018-12-21

DOI
10.1021/acssynbio.8b00400

Peer reviewed

Published in final edited form as:

ACS Synth Biol. 2018 December 21; 7(12): 2930–2934. doi:10.1021/acssynbio.8b00400.

Tunable expression systems for orthogonal DNA replication

Ziwei Zhong[1], **Arjun Ravikumar**[1], **Chang C. Liu**[1,2,3,*]

[1]Department of Biomedical Engineering, University of California, Irvine, California 92697, United States

[2]Department of Chemistry, University of California, Irvine, California 92697, United States

[3]Department of Molecular Biology & Biochemistry, University of California, Irvine, California 92697, United States

Abstract

We recently developed an orthogonal DNA replication (OrthoRep) system capable of driving the rapid continuous evolution of genes *in vivo*. However, OrthoRep uses a special transcription system whose components (*e.g.* promoters) have previously limited the strength with which OrthoRep-encoded genes can be expressed. Here, we report a collection of synthetic or evolved OrthoRep expression parts that allow OrthoRep-encoded genes to span expression levels matching those of endogenous *Saccharomyces cerevisiae* genes. Specifically, we found that various promoter mutations as well as a genetically-encoded poly(A) tail enable us to tune the expression level of OrthoRep-encoded genes over a large range and up to levels 43-fold higher than were previously attained, reaching at least ~40% of the strength of the genomic *TDH3* promoter. We further show that expression level gains using our new parts are stable over passaging and consistent across multiple genes and OrthoRep systems of different mutation rates. This new set of expression parts further expands OrthoRep's applicability to the continuous *in vivo* evolution of proteins and pathways.

Introduction

We have previously described an orthogonal replication system in *Saccharomyces cerevisiae*, OrthoRep, which allows for the *in vivo* directed evolution of genes in a continuous and rapid manner.[1,2] OrthoRep is an engineered derivative of the linear killer cytoplasmic plasmids from *Kluyveromyces lactis*, pGKL1 (p1) and pGKL2 (p2).[3] Since each of the plasmids encodes its own dedicated DNA polymerase (DNAP) and replicates

*Corresponding Author ccl@uci.edu.

Author Contributions Z.Z., A.R., and C.C.L. designed the experiments. Z.Z. and A.R. conducted the experiments. Z.Z., A.R., and C.C.L. analyzed the data and wrote the manuscript.

Associated Content

Supporting Information

Methods

Figure S1: Schematic of integration cassettes used in this study

Figure S2: Individual testing of K2O10 UCR mutants

Figure S3: OrthoRep expression using two different fluorescent reporters

Figure S4: Representative gating strategy for flow cytometry

Table S1: List of UCRs used in this study

independently of the host genome (Figure 1a),[2–4] the resulting property of replicative orthogonality has allowed us to engineer highly-error prone versions of p1's DNAP (TP-DNAP1). As a result, genes encoded on p1 continuously mutate ~100,000 times more rapidly than genes encoded on the genome while chromosomal genes remain at their natural mutation rate of 10^{-10} substitutions per base (s.p.b.). Consequently, genes encoded on OrthoRep rapidly evolve under selection, resulting in a scalable system for continuous *in vivo* evolution. Indeed, using OrthoRep, our group has mapped drug-target-based resistance in scores of replicates to elucidate new adaptive trajectories and their interplay[1] and are currently carrying out several other rapid evolution experiments.

Genes encoded on OrthoRep need to be expressed for selection to act on their function, and this occurs through transcription by a dedicated set of enzymes encoded on p2. These enzymes, which include an RNA polymerase (RNAP) and a capping enzyme, appear distinct and isolated from nuclear transcription. Nuclear promoters are nonfunctional on p1,[5] and p1/p2 promoter sequences fail to yield proteins when encoded on nuclear plasmids.[6] Additionally, OrthoRep transcripts lack canonical 3' mRNA polyadenylation, as northern and RACE analysis of p2 transcripts show they are similar in size to their encoded gene,[7,8] suggesting that host mRNA processing does not act on OrthoRep transcripts. Therefore, special cis-elements need to be used when encoding genes for OrthoRep.

To drive transcription, our previous experiments with OrthoRep have relied on the addition of upstream control regions (UCRs) to any heterologous gene of interest (GOI) that we wish to continuously evolve. These UCRs are simply defined as the 100 bp sequence upstream of the start codon of endogenous open reading frames (ORFs) encoded on unmodified p1 and p2 plasmids.[1,2,9] We found that the strengths of the promoters in these natural UCRs are relatively low. Indeed, the strongest promoter identified, derived from pG**K**L**2 O**RF**10** (K2O10), only drives expression, summed across the many copies of p1, at a level equivalent to a low-medium strength nuclear promoter.[2,9] While these promoters may allow for evolution of proteins whose initial selectable function does not require high expression to achieve (*e.g.* transcription factors, biosensors, and certain enzymes), we sought to expand the range of expression levels available to OrthoRep to make it more versatile in the evolution of proteins or pathways. By identifying promoter mutants from our previous OrthoRep-based continuous evolution experiments and combining these promoters with a genetically encoded 3' poly(A) tail, we now report a panel of expression cassettes that span the range of expression strengths achievable on standard yeast nuclear plasmids.

Results

UCR mutations increase OrthoRep expression

In many OrthoRep evolution experiments, the repeated occurrence of mutations in the UCR regions attached to the GOI undergoing evolution led us to hypothesize that these mutations increased expression of target genes. For example, in a replicate evolution experiment for DHFR-dependent pyrimethamine resistance, we found several mutations in DHFR's K2O10 UCR,[1] and it is known that DHFR overexpression is a mechanism for increasing resistance.[10] Using a fluorescent protein reporter as a proxy for expression strength, we individually verified six different substitution mutations in K2O10 UCR, and observed that they indeed

significantly increase fluorescence relative to wild type (wt) K2O10 UCR (Figure S2a,b, Table S1). These individual mutations were strategically combined to form a non-exhaustive panel of double, triple, and quadruple K2O10 UCR mutants for further analysis (Figure 1b, S2c,d, Table S1). One set of mutations, referred to as **10B2**, resulted in an up to three-fold increase in expression over the wt K2O10 UCR sequence (Figure 1b, S2c,d, Table S1).

Genetically encoded poly(A) tails boost expression of genes encoded in OrthoRep

Due to the lack of 3' polyadenylation on native OrthoRep transcripts, we hypothesized that the synthetic addition of 3' poly(A) tails would improve translation efficiency. Poly(A) tails are critical determinants of mRNA stability and are important for ribosome cycling, both of which contribute to protein expression.[11] In canonical eukaryotic transcription, poly(A) polymerases coordinate *in cis* with RNA polymerase II to tail nascent mRNA transcripts.[12] Therefore, nuclear poly(A) polymerases cannot be simply repurposed for OrthoRep. Instead, Dower and colleagues reported an alternative, synthetic route for polyadenylation wherein a genetically encoded poly(A) tract and self-cleaving hammerhead ribozyme (RZ) was positioned directly downstream of a GOI and transcribed by an RNA polymerase to mimic a canonical poly(A) tail.[13] We sought to implement this synthetic polyadenylation strategy for OrthoRep.

A genetically encoded poly(A) tail containing at least 48 adenosines significantly increased expression of GOIs encoded on p1 by up to an order of magnitude (Figure 1d). Consistent with previous reports that natural yeast mRNAs have a poly(A) length of ~75 nucleotides, [14,15] maximal protein expression was seen with the synthetic 75 adenosine tail (A75-RZ). As expected, the identity of nucleotides play a role in this expression level increase, as a similarly long tract of thymidines fails to achieve a significant increase. As a result, A75-RZ was used for all subsequent experiments. In particular, we combined A75-RZ with the various promoter mutants identified above, and found a consistent 10- to 20-fold increase in protein expression owing to A75-RZ. This was observed across different UCRs and independent experiments (Figure 1e), suggesting that promoter strength and transcript poly-adenylation independently affect protein expression, resulting in two means by which expression of GOIs can be tuned.

Promoters and poly(A) tails can be combined to yield a panel of genetic constructs for customized and high expression levels from OrthoRep

By combining different promoter mutants with A75-RZ, we created a panel of OrthoRep expression constructs that span the ~280-fold range achieved with standard constitutive genomic promoters on CEN/ARS plasmids (Figure 2a). We identified three promoters, K2O1, K2O10, and the evolved promoter **10B2**, which allow for custom-tailored expression of genes encoded on OrthoRep. In fact, **10B2** used in combination with A75-RZ drives gene expression at a minimum of 39% (95% CI [27.6, 53.6]) of the level of a strong pTDH3 nuclear promoter when summed across the copy numbers of p1 (~100)[1,2] and the CEN/ARS plasmid (~3)[16] on which pTDH3 was tested (Figure 2a, S3a). To ensure that this panel of promoters would be generalizable and stable for experimental evolution, we validated several promoters with a second fluorescent protein (Figure S3) and confirmed stability over at least 60 generations (Figure 2b).

10B2 and the poly(A) tail also increase expression levels over previous expression systems in the presence of TP-DNAP1–4-2, a highly error-prone DNAP.[1] Error-prone DNAPs are supplied *in trans* during evolution to introduce mutations onto genes encoded on OrthoRep (Figure S1b). However, the activity of these engineered error-prone DNAPs is lower than the wt TP-DNAP1, so they sustain lower copy numbers of p1, and consequently, lower protein expression from p1. For example, the copy number of p1 when replicated by TP-DNAP1–4-2 is ~6.[1] We tested our panel of promoters with error-prone TP-DNAP1–4-2 supplied *in trans* to characterize expression levels as well as to validate the panel of promoters for use during rapid evolution. Although the absolute levels of fluorescence were reduced in accordance with the previously seen copy number decrease, expression levels were still able to span over an 18-fold range and achieve at least 68% (95% CI [50.0, 92.6])) the levels of a medium strength pRPL18B nuclear promoter encoded on a standard CEN/ARS plasmid (Figure 2a, S3b). Since p1 copy number under replication by TP-DNAP1–4-2 is similar to the copy number of a CEN/ARS plasmid, this suggests that the per-copy expression strength for GOIs on OrthoRep can reach pRPL18B levels, which is much higher relative to our previous expression systems for OrthoRep and should be sufficient to achieve selectable function from most GOIs one might wish to evolve.

Discussion and Conclusion

We have engineered a panel of OrthoRep expression cassettes that allows expression levels of OrthoRep-encoded GOIs to span a range comparable to traditional yeast promoters on nuclear plasmids. Interestingly, the mutations that were able to combine to form **10B2** are 24 or 27 bases upstream of the start codon and are adjacent to the conserved ATNTGA sequence (Table S1, highlighted in yellow).[9] This suggests that these mutations may directly influence RNAP recruitment or initiation and that further optimization of RNAP recruitment may be possible as native UCRs are not optimized for high expression. Already, these improvements in expression should extend OrthoRep's application space to include some classic categories of directed evolution experiments that have been conducted using high expression promoters, including yeast surface display[17] and pathway engineering.[18] One consideration when choosing an expression cassette for OrthoRep is the effect of polymerase choice on cumulative expression. Our previous characterization of OrthoRep polymerases has shown that highly error-prone DNAPs reduce p1 copy number ~20 fold, with aggregate expression of GOIs decreasing comparably. However, since p2 remains unaffected by the presence of error-prone TP-DNAP1s,[4] the per-copy expression of GOIs encoded on OrthoRep remain unchanged, thus allowing beneficial mutations on GOIs both to achieve homoplasmy more rapidly as well as to have a proportionally higher impact on fitness. Further, we note that since OrthoRep is a continuous evolution system, the utility of these strong expression systems is mostly at the beginning of an evolution experiment, where initial diversity in the GOI generated through neutral drift with OrthoRep must yield a variant with enough function for selection to act upon. After that, continuous evolution will take hold, and the function and expression of the GOI will further improve as necessary.

These expression cassettes also establish a foundation for using OrthoRep as an orthogonal transcription system. This property of transcriptional orthogonality has important implications for the use of yeast strains in protein production or the construction of synthetic

biological circuits. In particular, genomic promoters designated as constitutive are still often affected by cellular regulation, leading to less predictable behavior of synthetic genetic circuits using such parts. However, OrthoRep's orthogonal replication and transcription properties should be insulated from such cellular regulation to a greater degree. Therefore, outside of OrthoRep's use as a continuous evolution system, it may be useful in the establishment of "virtual machine" platforms for encoding synthetic genes and circuits in complex hosts.[19] Finally, OrthoRep's replicative and transcriptional orthogonality should enable future engineering efforts aimed at creating polymerization systems that use only non-natural substrates in all nucleic acid information transfer steps in a GOI's expression,[20] and eventually association with non-natural ribosomes[21] to ultimately realize a full orthogonal central dogma.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

This work was funded by the NIH (1DP2GM119163-01) and the NSF (MCB 1545158). The A75-RZ constructs used for the initial analysis were generously provided by Dr. Michael Rosbash.

Abbreviations

DNAP	DNA polymerase
FP	fluorescent protein
GOI	gene of interest
RNAP	RNA polymerase
RZ	ribozyme
UCR	upstream control region
wt	wild type

References

(1). Ravikumar A, Arzumanyan GA, Obadi MKA, Javanpour AA, and Liu CC (2018) Scalable continuous evolution of genes at mutation rates above genomic error thresholds. Cell in press.

(2). Ravikumar A, Arrieta A, and Liu CC (2014) An orthogonal DNA replication system in yeast. Nat. Chem. Biol 10, 175–7. [PubMed: 24487693]

(3). Stark MJR, Boyd A, Milehams AJ, and Romanosg ANDMA (1990) The Plasmid-encoded Killer System of Kluyveromyces lactis : A Review.

(4). Arzumanyan GA, Gabriel KN, Ravikumar A, Javanpour AA, and Liu CC (2018) Mutually Orthogonal DNA Replication Systems in Vivo. ACS Synth. Biol 7, 1722–1729. [PubMed: 29969238]

(5). Kamper J, Meinhardt F, Gunge N, and Esser K (1989) In vivo construction of linear vectors based on killer plasmids from Kluyveromyces lactis: selection of a nuclear gene results in attachment of telomeres. Mol Cell Biol 9, 3931–3937. [PubMed: 2779572]

(6). Stam JC, Kwakman J, Meijer M, Stuitje AR, Cytology M, and Centre B (1986) Efficient isolation of the linear DNA killer plasmid of Kluyveromyces lactis: evidence for location and expression in the cytoplasm and characterization of the terminally bound proteins. Nucleic Acids Res. 14, 6871–6884. [PubMed: 3763395]

(7). Schaffrath R, Soond SM, and Meacock PA (1995) The DNA and RNA polymerase genes of yeast plasmid pGKL2 are essential loci for plasmid integrity and maintenance. Microbiology 141, 2591–2599. [PubMed: 7582020]

(8). Vopalensky V, Sykora M, Masek T, and Pospisek M (2018) Messenger RNAs transcribed from yeast linear cytoplasmic plasmids possess unconventional 5’ and 3’ UTRs and suggest a novel mechanism of translation. BioRxiv.

(9). Schickel J, Helmig C, and Meinhardt F (1996) Kluyveromyces lactis killer system: Analysis of cytoplasmic promoters of the linear plasmids. Nucleic Acids Res. 24, 1879–1886. [PubMed: 8657569]

(10). Thaithong S, Chan SW, Songsomboon S, Wilairat P, Seesod N, Sueblinwong T, Goman M, Ridley R, and Beale G (1992) Pyrimethamine resistant mutations in Plasmodium falciparum. Mol. Biochem. Parasitol 52, 149–157. [PubMed: 1620155]

(11). Sachs A (1990) The role of poly(A) in the translation and stability of mRNA. Curr. Opin. Cell Biol 2, 1092–1098. [PubMed: 2099802]

(12). Runner VM, Podolny V, and Buratowski S (2008) The Rpb4 subunit of RNA polymerase II contributes to cotranscriptional recruitment of 3’ processing factors. Mol. Cell. Biol 28, 1883–1891. [PubMed: 18195044]

(13). Dower K, Kuperwasser N, Merrikh H, and Rosbash M (2004) A synthetic A tail rescues yeast nuclear accumulation of a ribozyme-terminated transcript. RNA 10, 1888–1899. [PubMed: 15547135]

(14). Brown CE, and Sachs AB (1998) Poly(A) tail length control in Saccharomyces cerevisiae occurs by message-specific deadenylation. Mol. Cell. Biol 18, 6548–59. [PubMed: 9774670]

(15). Eckmann CR, Rammelt C, and Wahle E (2011) Control of poly(A) tail length. Wiley Interdiscip. Rev. RNA 2, 348–361. [PubMed: 21957022]

(16). Karim AS, Curran KA, and Alper HS (2013) Characterization of plasmid burden and copy number in Saccharomyces cerevisiae for optimization of metabolic engineering applications. FEMS Yeast Res 13, 107–116. [PubMed: 23107142]

(17). Boder ET, and Wittrup KD (1997) Yeast Surface Display for Screening Combinatorial Polypeptide Libraries. Nat. Biotechnol 15, 553–557. [PubMed: 9181578]

(18). Lee SM, Jellison T, and Alper HS (2012) Directed evolution of xylose isomerase for improved xylose catabolism and fermentation in the yeast Saccharomyces cerevisiae. Appl. Environ. Microbiol 78, 5708–5716. [PubMed: 22685138]

(19). Liu CC, Jewett MC, Chin JW, and Voigt CA (2018) Toward an orthogonal central dogma. Nat. Chem. Biol 14, 103–106. [PubMed: 29337969]

(20). Pinheiro VB, Taylor AI, Cozens C, Abramov M, Renders M, Zhang S, Chaput JC, Wengel J, Peak-Chew S-Y, McLaughlin SH, Herdewijn P, and Holliger P (2012) Synthetic Genetic Polymers Capable of Heredity and Evolution. Science (80-.). 336, 341–344.

(21). Orelle C, Carlson ED, Szal T, Florin T, Jewett MC, and Mankin AS (2015) Protein synthesis by ribosomes with tethered subunits. Nature 524, 119–124. [PubMed: 26222032]

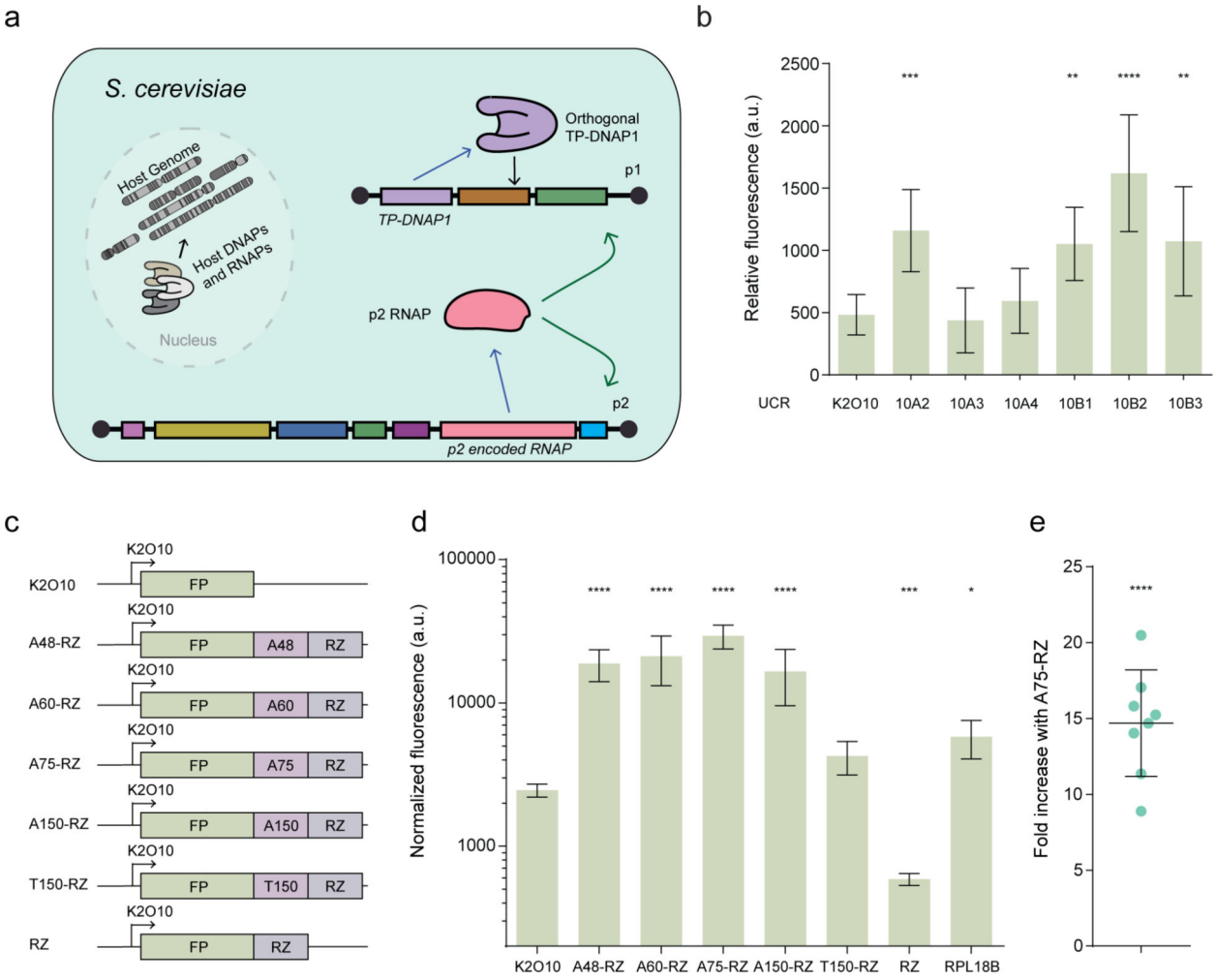

Figure 1.

A genetically encoded poly(A) tail increases expression of genes encoded on OrthoRep. (**a**) A schematic of OrthoRep, an orthogonal transcription/translation system in *Saccharyomyces cerevisiae*. The natural p1 plasmid encodes TP-DNAP1 for its own replication and p2 encodes the RNA polymerase for transcribing genes encoded on both p1 and p2. **(b)** Relative fluorescence of single mutants (10A prefix) or combination mutations (10B prefix) in K2O10 UCR. Statistical comparisons are vs. wt K2O10 by one-way ANOVA. (**c**) A graphical depiction of the constructs used to assay a genetically encoded 3' poly(A) tail. K2O10, the UCR of p2ORF10; FP, a fluorescent protein used as a reporter; A48, A60, etc., a tract of 48 adenosines, 60 adenosines, etc.; RZ, a 5' self-cleaving hammerhead ribozyme. (**d**) OD_{600}-normalized fluorescence of constructs shown in **c**. RPL18B denotes the use of pRPL18B on a CEN/ARS plasmid as a control for nuclear expression. N = 2–4. Statistical comparisons are vs. wt K2O10 by one-way ANOVA (**e**) Ratios of fluorescence with and without a poly(A) tail across 8 independent expression cassettes and experiments. Statistical comparison is vs. no difference by one-way t-test. Bars and errors denote mean and standard deviation or range. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$.

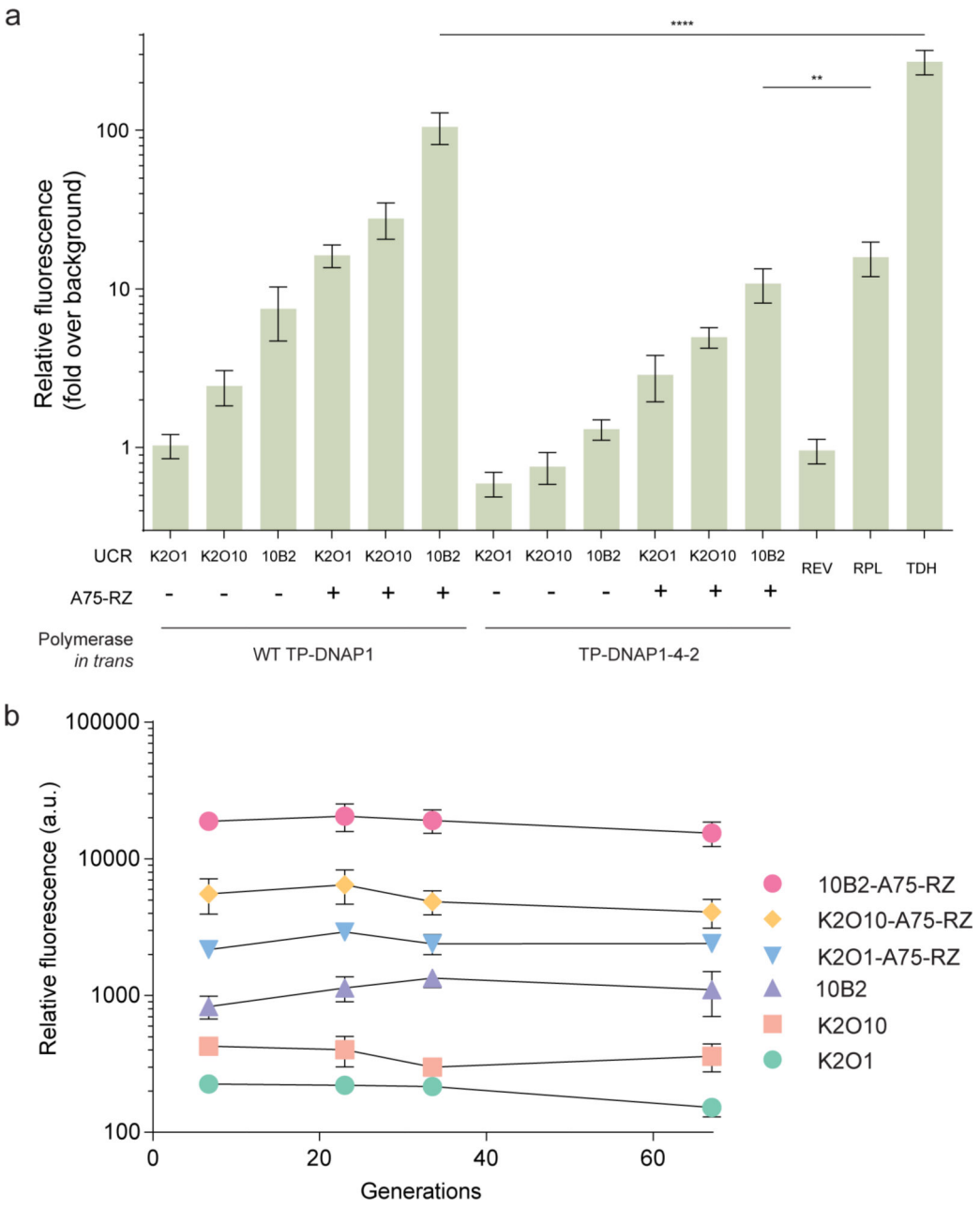

Figure 2.

A collection of OrthoRep expression cassettes spans a range of expression comparable to nuclear promoters. (**a**) Fluorescence fold change over background autofluorescence of three OrthoRep UCRs (K2O1, K2O10, and **10B2**) with or without a genetically encoded poly(A) tract and ribozyme (A75-RZ) in the context of wild-type (WT TP-DNAP1) or an error-prone polymerase (TP-DNAP1–4-2). REV, RPL, and TDH denote the use of pREV1, pRPL18B, or pTDH3 on nuclear CEN/ARS plasmids to drive the fluorescent reporter to serve as expression controls. Select comparisons are shown. ** $p < 0.01$, **** $p < 0.0001$ by one-

way ANOVA. (**b**) Relative fluorescence measurements of OrthoRep constructs over 67 generations. N = 6–8. Bars and errors denote mean and standard deviation.