

Supplemental Information

Creation of a eukaryotic multiplexed site-specific inversion system and its application for metabolic engineering

Jieyi Li^{#1,2,3}, Simiao Gong^{#1,2,3}, Yuan Ma^{#1,2,3}, Peiyan Han^{1,2,3}, Nan Wang^{1,2,3}, Zongheng Fu^{1,2,3}, Xinyi Zhang^{1,2,3}, Xinyang Huang^{1,2,3}, Tianyu Yang^{1,2,3}, Hanze Tong^{1,2,3}, Guang-Rong Zhao^{1,2,3}, Yi Wu^{1,2,3,*} and Ying-Jin Yuan^{1,2,3}

1. State Key Laboratory of Synthetic Biology, Tianjin University, Tianjin 300072, China;
2. Frontiers Science Center for Synthetic Biology and Key Laboratory of Systems Bioengineering (Ministry of Education), School of Chemical Engineering and Technology, Tianjin University, Tianjin 300072, China;
3. Frontiers Research Institute for Synthetic Biology, Tianjin University, Tianjin 300072, China.

* To whom correspondence should be addressed. Email: yi.wu@tju.edu.cn

This PDF file includes:

Supplemental Figure 1: Testing multisite inversion for the wild-type Rci and mutant Rci8.

Supplemental Figure 2: Third round of directed evolution of Rci in *Saccharomyces cerevisiae*.

Supplemental Figure 3: Testing the efficiency of inversion for Rci51-5 at different *sfx* sites.

Supplemental Figure 4: The multiplexed site-specific inversion system of Rci51-5/multi-*sfxa101* works in a mammalian cell.

Supplemental Figure 5: Validating the orthogonality of the Rci51-5/*sfxa101* system with other site-specific recombination systems.

Supplemental Figure 6: Design of the “AND gate” switch for precise control of the Rci51-5.

Supplemental Figure 7: The intergration of transcriptional unit for Rci51-5 into the genome.

Supplemental Figure 8: The proportion of dark orange colonies to total colonies at different induction times.

Supplemental Figure 9: The long-term stability analysis of the multi-gene promoter regulation system for β -carotene.

Supplemental Figure 10: The multi-gene promoter regulation system for violacein.

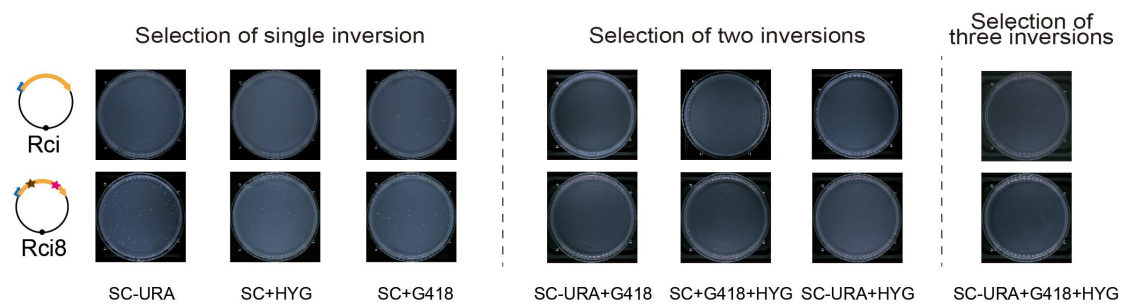
Supplemental Figure 11: The frequency of promoter substitution for the exogenous genes.

Supplemental Table 1: Strains used in this study.

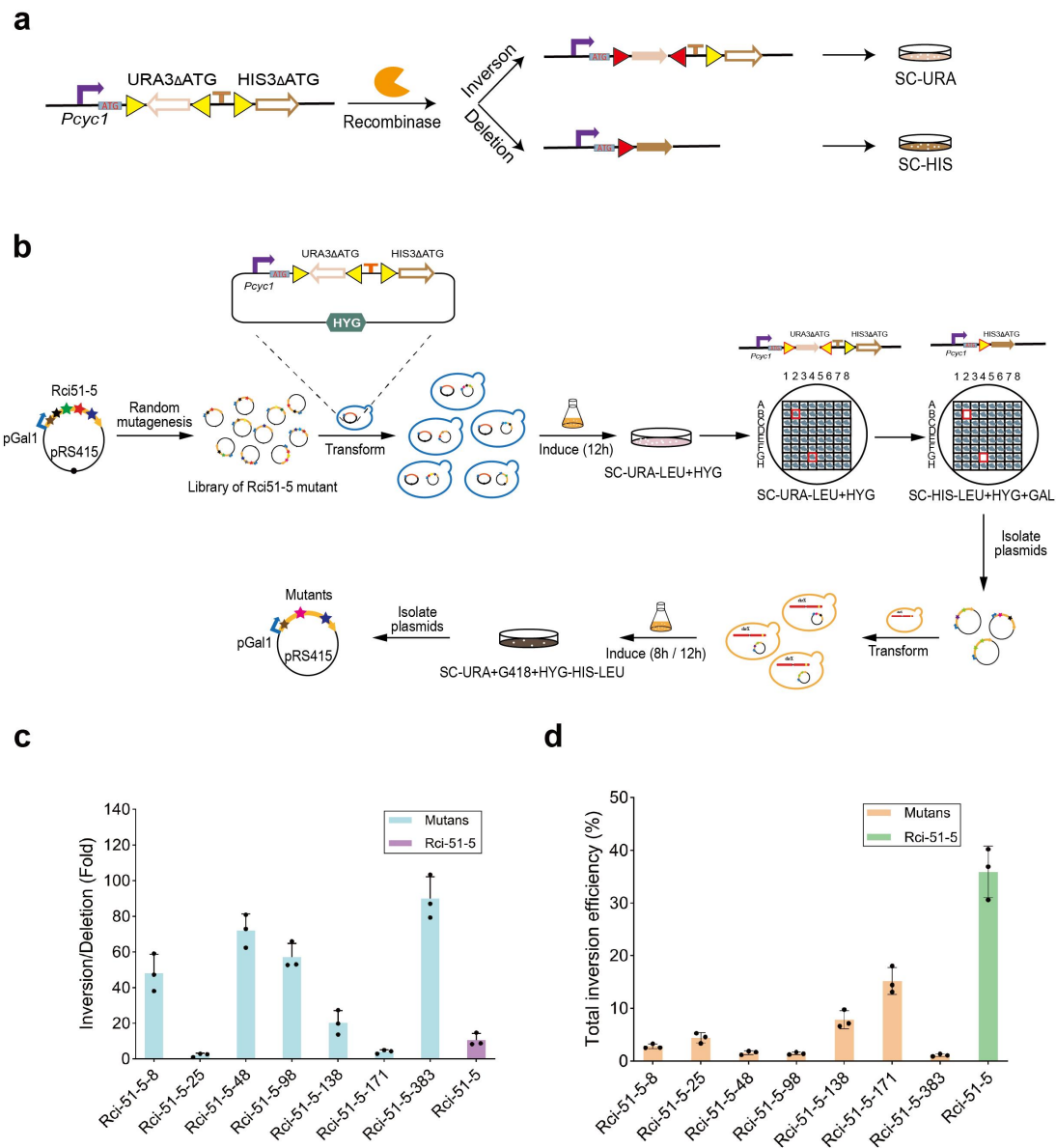
Supplemental Table 2: Plasmids used in this study

Supplemental Table 3: Primers used in this study for junction PCR

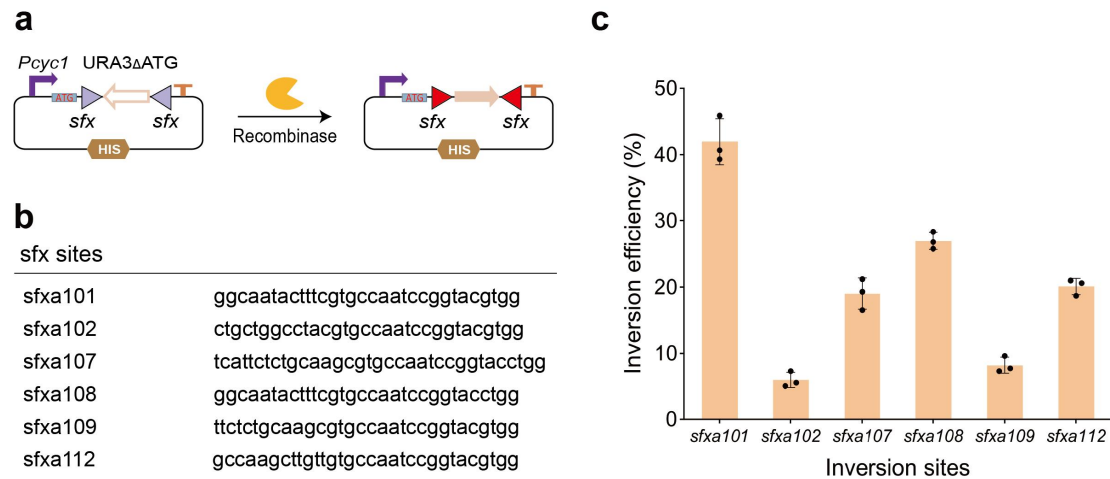
The sequences of genes involved in this study



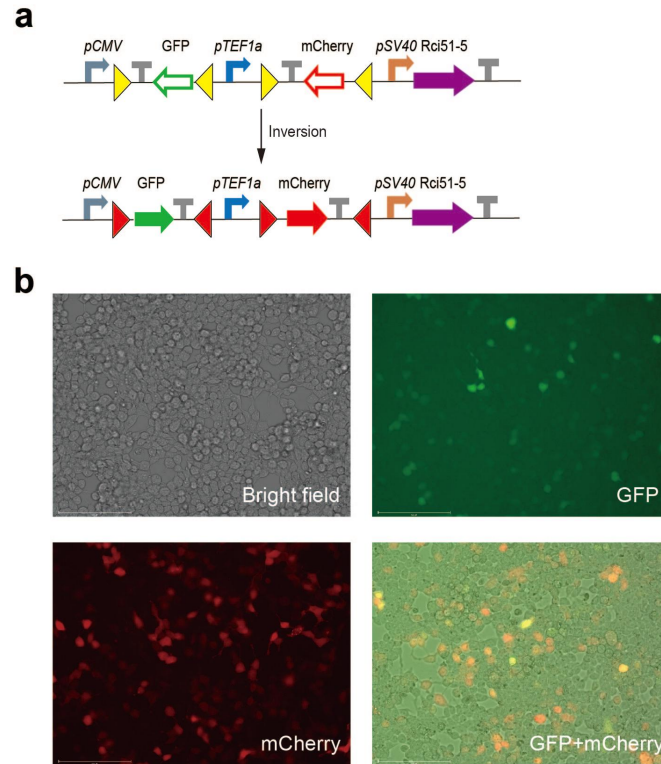
Supplemental Figure 1: Testing multisite inversion for the wild-type Rci and mutant Rci8. Yeast cultures containing the wild-type Rci or Rci8 were plated on various selective mediums after inducing 12 h, respectively. For both mutants, colonies were only observed on mediums of a single selection. The pictures were taken after 3 days incubation at 30°C.



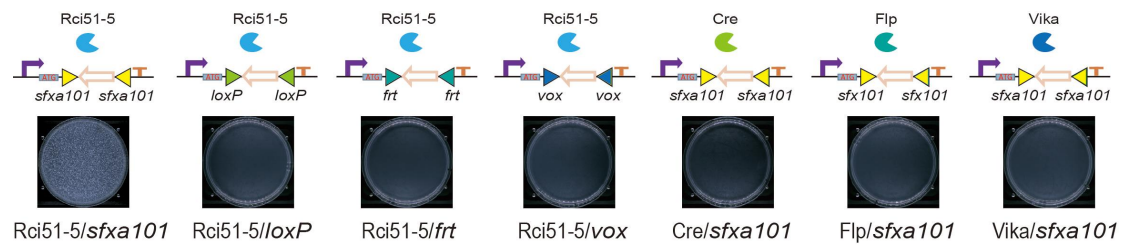
Supplemental Figure 2: Third round of directed evolution of Rci in *Saccharomyces cerevisiae*. **a.** The screening system for inversion enzymes of specificity in the third round of directed evolution. ATG* represents ATGAA. Two AA base pairs are inserted after ATG to avoid heterogeneous mutations caused by 31 bp *sfxa101* sites. **b.** Flow diagram of the third round of directed evolution of multiplexed site-specific inversion enzyme in *Saccharomyces cerevisiae*. **c.** The ratio of inversion/deletion frequency for varied mutants of Rci51-5. The blue represents the mutants generated in the third round, and the control is the template Rci51-5 (purple). The system for verification is shown in Supplemental Figure S3a. The error bars represent standard deviation of three biological replicates. **d.** The total inversion efficiency of mutants generated from the third round of directed evolution. The inversion efficiency of the mutants (orange) is much lower than the Rci51-5 (green). The error bars represent standard deviation of three biological replicates. Source data for this figure are provided as a Source Data file.



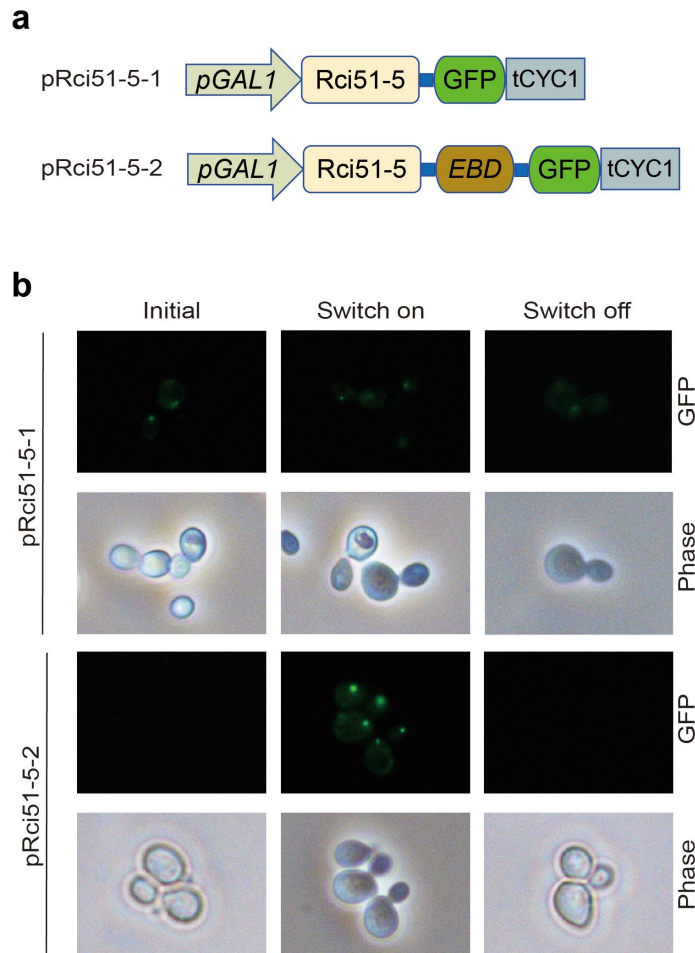
Supplemental Figure 3: Testing the efficiency of inversion for Rci51-5 at different *sfx* sites. **a.** Scheme of the recombination assay. The triangle (purple) represents the *sfx* sites. The red triangles indicate the sites that mediate recombination. **b.** Sequences of different *sfx* sites. **c.** The inversion efficiency of the mutant Rci51-5 at different *sfx* sites. The error bars represent standard deviation of three biological replicates. Source data for this figure are provided as a Source Data file.



Supplemental Figure 4: The multiplexed site-specific inversion system of Rci51-5/multi-*sfxa101* works in a mammalian cell. **a.** Schematic diagram of the inversion screening system in a mammalian cell. **b.** Representative microscopic images of color conversion of 293T cells after expression of Rci51-5. The Fluorescence microscopy of 293T cells was captured after 48 h post transfection. Scale bars, 125 μm . Source data for this figure are provided as a Source Data file.

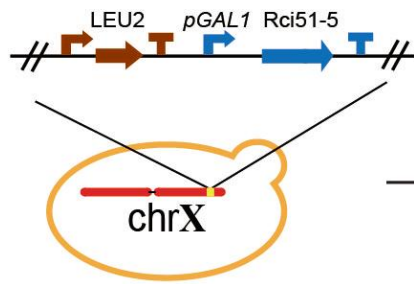


Supplemental Figure 5: Validating the orthogonality of the Rci51-5/*sfxa101* system with other site-specific recombination systems. The method of verification is the same as the Supplemental Figure 4a. The pictures were taken after 3 days incubation at 30°C.

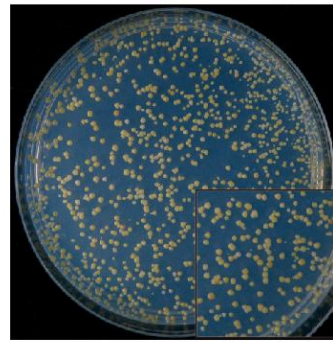


Supplemental Figure 6: Design of the “AND gate” switch for precise control of the Rci51-5. **a.** Design of the AND gate for inversion enzyme Rci51-5 (pRci51-5-2) and the control. Activation of the “AND gate” pRci51-5-2 required the input of galactose and estradiol simultaneously. The control pRci51-5-1 was regulated by galactose inducible promoter *GALI*. **b.** The “switch on” media for pRci51-5-1 and pRci51-5-2 were SGal-Leu medium and SGal-Leu medium containing 1 μ M estradiol, respectively. The “initial state” and the “switch off” media were SC-Leu medium. The GFP fluorescence of the cells were imaged from this cell lawn using a fluorescence microscope (Nikon Y-TV55). Scale bars, 10 μ m.

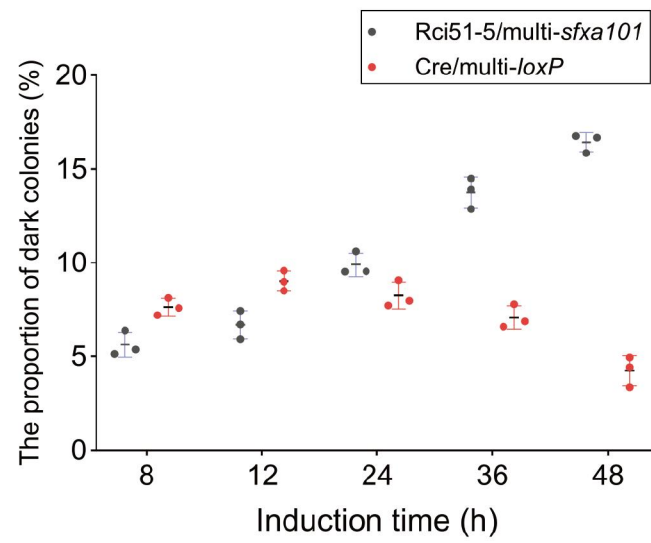
Integrate into the **X** chromosome



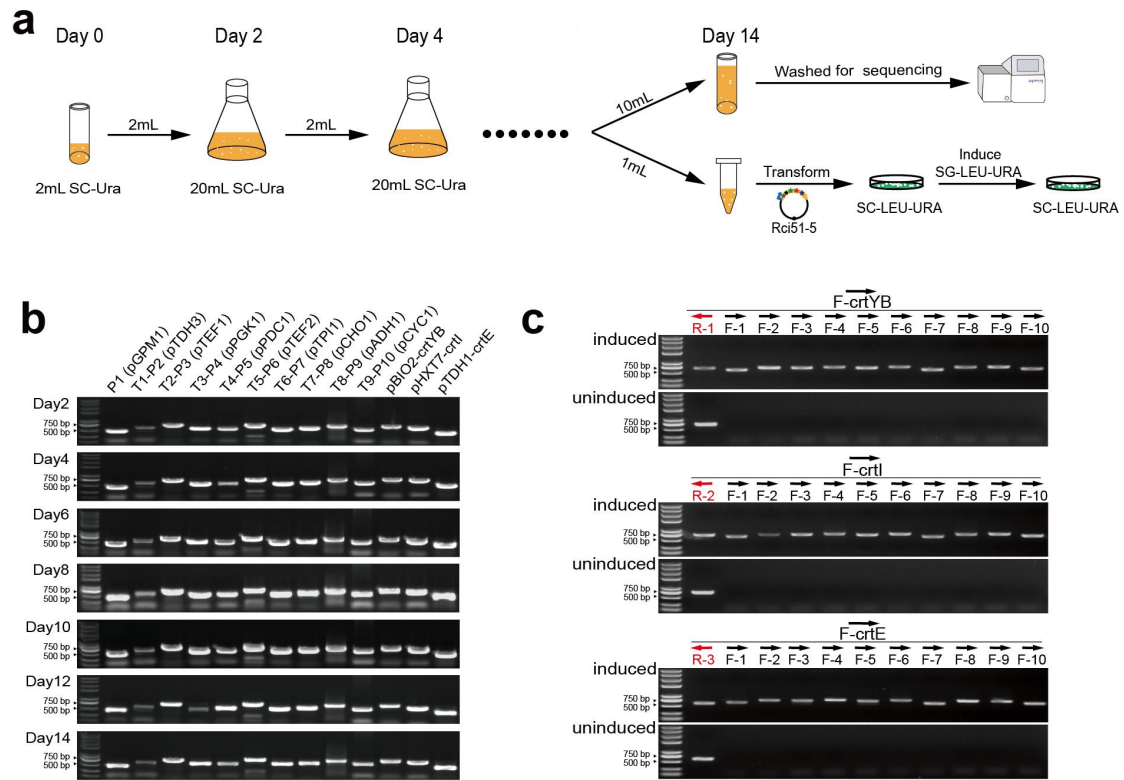
Induce



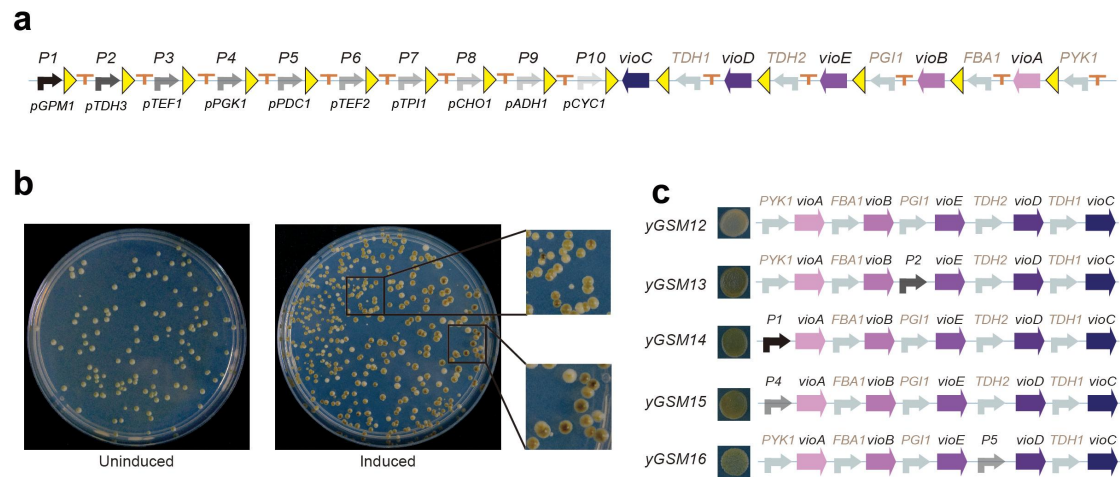
Supplemental Figure 7: The intergration of transcriptional unit for Rci51-5 into the genome.
Source data for this figure are provided as a Source Data file.



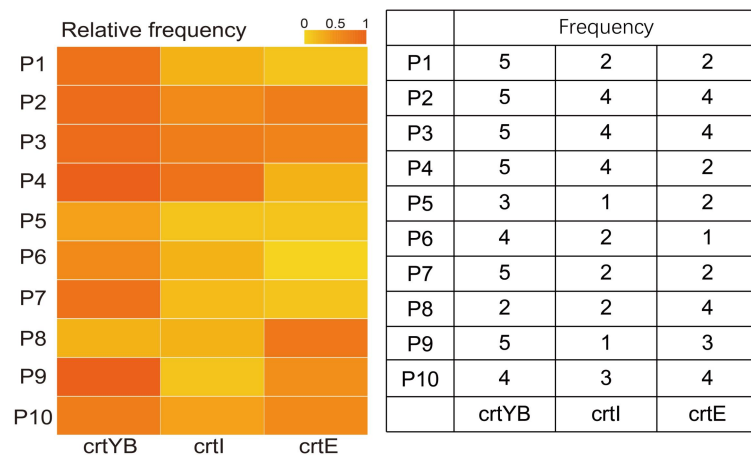
Supplemental Figure 8: The proportion of dark orange colonies to total colonies at different induction times. Bars and error bars represent average and standard deviation of three biological replicates, respectively. Source data for this figure are provided as a Source Data file.



Supplemental Figure 9: The long-term stability analysis of the multi-gene promoter regulation system for β -carotene. **a.** Schematic diagram of the passage and strategy of analysis. The yeast strain containing the system (*yLJY023*) was subjected to six serial passages for 14 days of growth in the presence of auxotroph screening (SC-Ura medium) at 30° C, 220rpm. **b.** PCRTag analysis of propagation stability of the system in *S. cerevisiae* by yeast colony PCR. Individual colonies were randomly selected from solid agar plates in each passage to validate the sequence integrity with 13 primer pairs. **c.** Three sets of PCR results for the pooled cells after induction (the similar was mentioned in Fig. 4b). Source data for this figure are provided as a Source Data file.



Supplemental Figure 10: The multi-gene promoter regulation system for violacein. **a.** The circuit of multi-gene promoter regulation system for violacein. The 3' and 5' of the circuit consist of the promoters array and exogenous genes, respectively. The yellow triangles represent the *sfxa101* site. **b.** Yeast colonies induced by this system. The uninduced serves as a control. The magnified region shows different colony colors after induction. **c.** The results for exogenous gene promoter substitution in 4 strains with dark colors are verified by PCR analysis. The control *yGSM12* is not induced. Source data for this figure are provided as a Source Data file.



Supplemental Figure 11: The frequency of promoter substitution for the exogenous genes. The figure on the left shows the relative frequency of promoter substitution and the highest frequency value is set to 1 for substitution. Different original promoters exhibit varying frequencies of substitution (color scale). The figure on the right shows the frequency of promoter substitution.

Supplemental Table 1: Strains used in this study

Strains	Description	Sources
BY4741	MATa, his3Δ1, leu2Δ0, met15Δ0, ura3Δ0	Our lab
yLJY001	BY4741 with High-throughput screening systems	This study
yLJY002	yLJY001 with pLJY001	This study
yLJY003	yLJY001 with pLJY002	This study
yLJY004	yLJY001 with pLJY003	This study
yLJY005	yLJY001 with pLJY004	This study
yLJY006	yLJY001 with pLJY005	This study
yLJY007	yLJY001 with pLJY006	This study
yLJY008	yLJY001 with pLJY007	This study
yLJY009	yLJY001 with pLJY008	This study
yLJY010	yLJY001 with pLJY009	This study
yLJY011	yLJY001 with pLJY010	This study
yLJY012	yLJY001 with pLJY011	This study
yLJY013	yLJY001 with pLJY012	This study
yLJY016	yLJY001 with pLJY015	This study
yLJY017	yLJY001 with pLJY016	This study
yLJY018	yLJY001 with pLJY017	This study
yLJY019	yLJY001 with pLJY018	This study
yLJY020	BY4741 with pLJY015	This study
yLJY022	yLJY020 with pLJY038	This study
yLJY023	yLJY020 with pLJY039	This study
yLJY024	yLJY023 after induced	This study
yLJY025	yLJY023 after induced	This study
yLJY028	yLJY023 after induced	This study
yLJY031	yLJY023 after induced	This study
yLJY033	yLJY023 after induced	This study
yLJY037	yLJY023 after induced	This study
yLJY044	yLJY023 after induced	This study
yLJY045	yLJY023 after induced	This study
yLJY049	yLJY023 after induced	This study
yLJY055	yLJY023 after induced	This study
yLJY062	yLJY023 after induced	This study
yLJY074	yLJY023 after induced	This study
yLJY081	yLJY023 after induced	This study
yGSM12	yLJY020 with pGSM063	This study
yGSM13	yGSM12 after induced	This study
yGSM14	yGSM12 after induced	This study
yGSM15	yGSM12 after induced	This study
yGSM16	yGSM12 after induced	This study

Supplemental Table 2: Plasmids used in this study

Plasmids	Description	Sources
pRS413	CEN/ARS with HIS3 marker	Our lab
pRS414	CEN/ARS with TRP1 marker	Our lab
pRS415	CEN/ARS with LEU2 marker	Our lab
pHPY001	pRS415-pGAL1-Rci	Our lab
pHPY004	pRS414-hphNT1-URA3-HIS3 with three sfxa101 sites	Our lab
pLJY001	Rci2 mutant	This study
pLJY002	Rci4 mutant	This study
pLJY003	Rci7 mutant	This study
pLJY004	Rci19 mutant	This study
pLJY005	Rci26 mutant	This study
pLJY006	Rci30 mutant	This study
pLJY007	Rci35 mutant	This study
pLJY008	Rci47 mutant	This study
pLJY009	Rci51 mutant	This study
pLJY010	Rci65 mutant	This study
pLJY011	Rci66 mutant	This study
pLJY012	Rci69 mutant	This study
pLJY015	Rci51-5 mutant	This study
pLJY017	Rci51-7 mutant	This study
pLJY018	Rci51-8 mutant	This study
pLJY019	Rci51-5-8 mutant	This study
pLJY020	Rci51-5-25 mutant	This study
pLJY021	Rci51-5-48 mutant	This study
pLJY022	Rci51-5-98 mutant	This study
pLJY023	Rci51-5-138 mutant	This study
pLJY024	Rci51-5-171 mutant	This study
pLJY025	Rci51-5-383 mutant	This study
pLJY026	recombination reporter vectors containing the sfx101 site	This study
pLJY027	recombination reporter vectors containing the sfx102 site	This study
pLJY028	recombination reporter vectors containing the sfx107 site	This study
pLJY029	recombination reporter vectors containing the sfx108 site	This study
pLJY030	recombination reporter vectors containing the sfx109 site	This study
pLJY031	recombination reporter vectors containing the sfx112 site	This study
pLJY032	recombination reporter vectors containing the loxP site	This study
pLJY033	recombination reporter vectors containing the vox site	This study
pLJY034	recombination reporter vectors containing the flp site	This study
pLJY035	Cre recombinase	This study
pLJY036	Flp recombinase	This study
pLJY037	Vika recombinase	This study
pLJY038	Multigene promoter regulation system for the β -carotene metabolic pathway by Cre/multi-loxP	This study

pLJY039	Multigene promoter regulation system for the β -carotene metabolic pathway by Rci51-5/multi-sfxa101	This study
pGSM063	Multigene promoter regulation system for the violacein metabolic pathway by Rci51-5/multi-sfxa101	This study

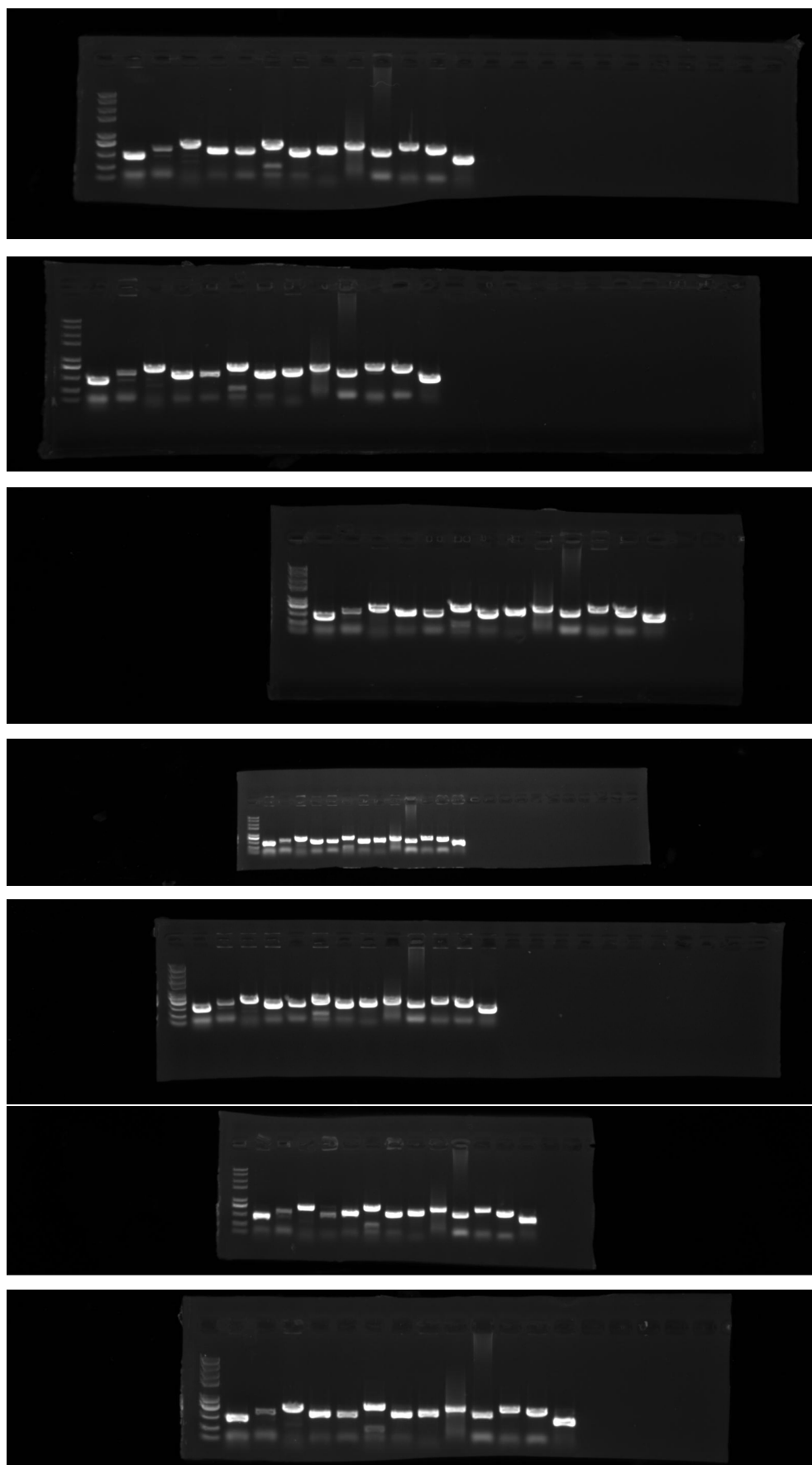
Supplemental Table 3: Primers used in this study for junction PCR

Primer name	Oligo Sequence (5'-3')
F-crtYB	aagactttcgtactctcctc
R-1	attgcagcaatcatgatccc
F-crtI	agaaagcaagaaccaacg
R-2	agatggattgatcggtcttg
F-crtE	tcacagaggatatcggctag
R-3	tacggaacatcctcacag
F-1	TAGCCATAGTGATGTCTAAGTAACC
F-2	TGTCTTTCGTTCAATCTCCAGG
F-3	TGCCAAGGCTTAAGTGAATTAC
F-4	ATATCTTAACAGGCGACAGTCC
F-5	TCTCTTCCCATCCTTTACG
F-6	ACCAACGCTAAGCAATAAGC
F-7	TTCTTGCCCATCAGATTGATG
F-8	ATGTTCTTGTAAGGGACGGTAG
F-9	AGTTTTTGTTTCATGCCTGGC
F-10	ACTCAAGTGATACGTGCATG

The sequences of gene involved in this study

> Rci sequence	ATGCCTTCACCTAGGATTCGTAATAATGTCCTGTCAAGGGCGCTAGATAAATATCTTAAACTGTGAGTGTCCATAAAAAGGGGCACC AACAGAATTCTATCGTTCAAATGTCATCAAAAGGTATCCTATTGCTTTAAGAAACATGGACGAGATAACCACGGTCGACATAGCCAC ATACCGTGATGTTAGATTGGCCGAAATCAACCCAAGAACAGGAAAACCTATTACTGGAACACTGTGAGATTGGAGCTGGCCCTGCT TTCATCCTGTTCAATATTGCAAGAGTCGAATGGGGTACGTGCAGGACTAACCCTGTTGAACTAGTTCGTAACCAAAAAGTGTCTAGT GGTAGGGATAGAAGGTTAACATCCAGTGAAGAAAGAAGACTGTCAAGATATTTAGAGAAAAGAACTTAATGTTATATGTTATTTCC ATTTGGCCTTAGAAAACAGCAATGAGACAAGGTGAAATATTAGCTTTGCGTTGGGAACATATAGATTTACGTACCGGTGTAGCCCATTT ACCTGAAACCAAAAATGGTCACAGCAGAGATGTTCCACTTAGTAGAAGAGCTCGTAATTTCTTGCAATGATGCCAGTCAATTTGCA CGGTAACGTATTTGACTACACCGCTTCCGGTTTTAAAAACGCATGGAGAATTGCCACTCAAAGACTTAGAATAGAAGACCTGCACCTC CATGACTTAAGACACAGGGCCATATCAAGATTTTGAATGGGTAGCTTGAACGTGATGGAATAGCTGCGATTAGTGGGCACAGA TCTATGAATATGTTAAAGCGTTACACTCACCTTAGAGCGTGGCAGTTAGTGAGCAAGTTAGACGCAAGAAGAAGGCAAACTCAGAA GGTCGCAGCTTGGTTTGACCTTATCCGGCACATATTACTACTATCGATGAAGAGAATGGTCAAAAAGCCCATAGAATAGAGATAGGT GATTTTGATAATTTGCATGTAACCTGCGACCACTAAAGAGGAAGCTGTACATAGAGCATCCGAGGTTCTATTGAGAACCTTAGCGATCG CTGCACAGAAGGGTGAAAGAGTTCCTTACCAGGAGCTTGGCCGTTAACGACCCGGATTATATTATGATCTGTCCCTTGAACCTGTG GTTCTACACCTTTATAA
>Rci8 sequence	ATGCCTTCACCTAGGATTCGTAATAATGTCCTGTCAAGGGCGCTAGATAAATATCTTAAACTGTGAGTGTCCATAAAAAGGGGCACC AACAGAATTCTATCGTTCAAATGTCATCAAAAGGTATCCTATTGCTTTAAGAAACATGGACGTGATAACCACGGTCGACATAGCCACA TACCGTGATGTTAGATTGGCCGAAATCAACCCAAGAACAGGAAAACCTATTACTGGAACACTGTGAGATTGGAGCTGGCCCTGCTT TCATCCTGTTCAATATTGCAAGAGTCGAATGGGGTACGTGCAGGACTAACCCTGTTGAACTAGTTCGTAACCAAAAAGTGTCTATGG TAGGGATAGAAGGTTAACATCCAGTGAAGAAAGAAGACTGTCAAGATATTTAGAGAAAAGAACTTAATGTTATATGTTATTTCCAT TTGGCCTTAGAAAACAGCAATGAGACAAGGTGAAATATTAGCTTTGCGTTGGGAACATATAGATTACGTACCGGTGTAGCCCATTTAC CTGAAACCAAAAATGGTCACAGCAGAGATGTTCCACTTAGTAGAAGAGCTCGTAATTTCTTGCAATGATGCCAGTCAATTTGCACG GTAACGTATTTGACTACACCGCTTCCGGTTTTAAAAACGCATGGAGAATTGCCACTCAAAGACTTAGAATAGAAGACCTGCACCTTCCA TGACTTAAGACACGAGGCCATATCAAGATTTTGAATGGGTAGCTTGAACGTGATGGAATAGCTGCGATTAGTGGGCACAGATC TATGAATATGTTAAAGCGTTACACTCACCTTAGAGCGTGGCAGTTAGTGAGCAAGTTAGACGCAAGAAGAAGGCAAACTCAGAAGG TCGCAGCTTGGTTTATACCTTATCCGGCACATATTACTACTATCGATGAAGAGAATGGTCAAAAAGCCCATAGAATAGAGATAGGTGAT TTTGATAATTTGCATGTAACCTGCGACCACTAAAGAGGAAGCTGTACATAGAGCATCCGAGGTTCTATTGAGAACCTTAGCGATCGCTG CACAGAAGGGTGAAAGAGTTCCTTACCAGGAGCTTGGCCGTTAACGACCCGGATTATATTATGATCTGTCCCTTGAACCTGTGTT TACACCTTTATAA
> Rci51 sequence	ATGCCTTCACCTAGGATTCGTAATAATGTCCTGTCAAGGGCGCTAGATAAATATCTTAAACTGTGAGTGTCCATAAAAAGGGGCACC AACAGAATTCTATCGTTCAAATGTCATCAAAAGGTATCCTATTGCTTTAAGAAACATGGACGAGATAACCACGGTCGACATAGCCAC ATACCGTGATGTTAGATTGGCCGAAATCAACCCAAGAACAGGAAAACCTATTACTGGAACACTGTGAGATTGGAGCTGGCCCTGCT TTCATCCTTGTCAATATTGCAAGAGTCGAATGGGGTACGTGCAGGACTAACCCTGTTGAACTAGTTCGTAACCAAAAAGTGTCTAGT GGTAGGGATAGAAGGTTAACATCCAGTGAAGAAAGAAGACTGTCAAGATATTTAGAGAAAAGAACTTAATGTTATATGTTATTTCC ATTTGGCCTTAGAAAACAGCAATGAGACAAGGTGAAATATTAGCTTTGCGTTGGGAACATATAGATTTACGTACCGGTGTAGCCCATTT ACCTGAAACCAAAAATGGTCACAGCAGAGATGTTCCACTTAGTAGAAGAGCTCGTAATTTCTTGCAATGATGCCAGTCAATTTGCA CGGTAACGTATTTGACTACACCGCTTCCGGTTTTAAAAACGCATGGAGAATTGCCACTCAAAGACTTAGAATAGAAGACCTGCACCTC CATGACTTAAGACACAGGGCCATATCAAGATTTGTTGAATGGGTAGCTTGAACGTGATGGAATAGCTGCGATTAGTGGGCACAGA TCTATGAATATGTTAAAGCGTTACACTCACCTTAGAGCGTGGCAGTTAGTGAGCAAGTTAGACGCAAGAAGAAGGCAAACTCAGAA GGTCGCATCTTGGTTTGACCTTATCCGGCACATATTACTACTATCGATGAAGAGAATGGTCAAAAAGCCCATAGAATAGAGATAGGTG ATTTTGATAATTTGCATGTAACCTACGACCACTAAAGAGGAAGCTGTACATAGAGCATCCGAGGTTCTATTGAGAACCTTAGCGATCGC TGACAGAAGGGTGAAAGAGTTCCTTACCAGGAGCTTGGCCGTTAATGACCCGGATTATATTATGATCTGTCCCTTGAACCTGTG TCTACACCTTTA
> Rci51-5 sequence	ATGCCTTCACCTAGGATTCGTAATAATGTCCTGTCAAGGGCGCTAGATAAATATCTTAAACTGTGAGTGTCCATAAAAAGGGGCACC AACAGAATTCTATCGTTCAAATGTCATCAAAAGGTATCCTATTGCTTTAAGAAACATGGACGAGATAACCACGGTCGACATAGCCAC ATACCGTGATGTTAGATTGGCCGAAATCAACCCAAGAACAGGAAAACCTATTACTGGAACACTGTGAGATTGGAGCTGGCCCTGCT TTCATCCTGTTCAATATTGCAAGAGTCGAATGGGGTACGTGCAGGACTAACCCTGTTGAACTAGTTCGTAACCAAAAAGTGTCTAGT GGTAGGGATAGAAGGTTAACATCCAGTGAAGAAAGAAGACTGTCAAGATATTTAGAGAAAAGAACTTAATGTTATATGTTATTTCC ATTTGGCCTTAGAAAACAGCAATGAGACAAGGTGAAATATTAGCTTTGCGTTGGGAACATATAGATTTACGTACCGGTGTAGCCCATTT ACCTGAAACCAAAAATGGTCACAGCAGAGATGTTCCACTTAGTAGAAGAGCTCGTAATTTCTTGCAATGATGCCAGTCAATTTGCA CGGTAACGTATTTGACTACACCGCTTCCGGTTTTAAAAACGCATGGAGAATTGCCACTCAAAGACTTAGAATAGAAGACCTGCACCTC CATGACTTAAGACACAGGGCCATATCAAGATTTGTTGAATGGGTAGCTTGAACGTGATGGAATAGCTGCGATTAGTGGGCACAGA TCTATGAATATGTTAAAGCGTTACACTCACCTTAGAGCGTGGCAGTTAGTGAGCAAGTTAGACGCAAGAAGAAGGCAAACTCAGAA GGTCGCATCTTGGTTTGACCTTATCCGGCACATATTACTACTATCGATGAAGAGAATGGTCAAAAAGCCCATAGAATAGAGATAGGTG ATTTTGATAATTTGCATGTAACCTACGACCACTAAAGAGGAAGCTGTACATAGAGCATCCGAGGTTCTATTGAGAACCTTAGCGATCGC TGACAGAAGGGTGAAAGAGTTCCTTACCAGGAGCTTGGCCGTTAATGACCCGGATTATATTATGATCTGTCCCTTGAACCTGTG TCTACACCTTTA

Uncropped scans of gels in Supplemental Figure 9b



Uncropped scans of gels in Supplemental Figure 9c

