Mutagenesis system by fusing cytidine deaminase with T7 promoter in yeast

Zi-rui Huang[1], Bing-Zhi Li[2], Xiang-rong Chen[1], Dan-feng Liu[1], You-zhi Cui[1], and Yingjin Yuan[3]

[1]Affiliation not available
[2]Tianjin University
[3]School of Chemical Engineering and Technology

March 27, 2023

Abstract

The occurrence of random mutations can increase the diversity of the genome and promote the evolutionary process of organisms. High efficiency mutagenesis techniques significantly accelerate the evolutionary process. In this work, we describe a targeted in vivo mutagenesis system to significantly increase mutation frequency and generate mutations across all four nucleotides. We constructed different DNA-modifying enzyme-PmCDA1-T7 RNA polymerase fusion proteins, achieved targeted mutagenesis by flanking the target gene with T7 promoters, and tuned the mutation spectra by introducing different DNA-modifying enzymes. With the mutagenesis fusion proteins, the mutation frequency of the target gene could reach 5.13x10-3, and the proportion of non-C-T mutations is 10˜11-fold higher than the cytidine-based evolutionary tools. We also demonstrated that our mutagenesis tools could be used to evolve the essential enzyme in the β-carotene biosynthesis process and generate mutations with different types.

Mutagenesis system by fusing cytidine deaminase with T7 promoter in yeast

Zi-Rui Huang, Xiang-Rong Chen, Dan-Feng Liu, You-Zhi Cui, Bing-Zhi Li*, Ying-Jin Yuan

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, PR China

*Corresponding author: Prof. Bing-Zhi Li (bzli@tju.edu.cn)

Abstract

The occurrence of random mutations can increase the diversity of the genome and promote the evolutionary process of organisms. High efficiency mutagenesis techniques significantly accelerate the evolutionary process. In this work, we describe a targeted *in vivo* mutagenesis system to significantly increase mutation frequency and generate mutations across all four nucleotides. We constructed different DNA-modifying enzyme-PmCDA1-T7 RNA polymerase fusion proteins, achieved targeted mutagenesis by flanking the target gene with T7 promoters, and tuned the mutation spectra by introducing different DNA-modifying enzymes. With the mutagenesis fusion proteins, the mutation frequency of the target gene could reach $5.13x10^{-3}$, and the proportion of non-C-T mutations is 10˜11-fold higher than the cytidine-based evolutionary tools. We also demonstrated that our mutagenesis tools could be used to evolve the *CrtE* , *CrtI* , and *CrtYB* genes in yeast to increase β-carotene yields.

Key words:

mutagenesis; cytidine deaminase fusion; synthetic biology; T7 promoter**;** *Saccharomyces cerevisiae*

Introduction

Mutations can increase genome diversity and drive the evolution of organisms. Random mutagenesis plays a crucial role in industrial strain breeding, new drug research and development, protein engineering and many other aspects.[1-6] However, under natural conditions, the genomic error rate is as low as 10^{-8}~10^{-9}, which makes natural evolution a very slow process. Therefore, how to increase the mutation rate of target genes, expand the mutation spectrum, and obtain the mutants with desired phenotypes efficiently are the essential problems that researchers have to consider.

There has been a lot of research into accelerating the evolution over the last few decades. Several strategies have been employed by researchers to accelerate the evolution of target genes, resulting in impressive progress. Radiation or chemically-induced mutagenesis is a conventional and commonly-used strategy to generate random mutations.[7-9] This method effectively increases the mutation rate of the target gene and is relatively straightforward to implement. As CRISPR/Cas techniques continue to advance, a multitude of CRISPR-based targeted mutagenesis systems have emerged, including CHAnGE, MAGESTIC, and many others.[10-13] In these systems, Cas9 variants are capable of precisely identifying the targeted gene, while gRNAs facilitate accurate and traceable editing of the gene. This leads to the production of numerous mutants with varying genotypes. By using CRISPR-based targeting systems, mutations in non-targeted regions are avoided. Additionally, traceable editing via barcoded gRNAs could provide valuable insights into the molecular mechanisms of evolution. To overcome the constraints of limited DNA library size and low DNA library-transformation efficiencies, some researchers have fused the Cas9-variants with error-prone polymerases or deaminases such as AID to create random mutations in the target regions.[14-20] Mutations are commonly located in an area adjacent to the target site spanning around 40 to 80 base pairs. In recent years, there has been a rise in the use of targeted in vivo mutagenesis tools such as cytidine/adenosine deaminase, TRACE,[21] eMutaT7,[22] and TRIDENT,[23] which are based on T7 RNA polymerase-deaminase fusion proteins.[24-27]Deaminases greatly increase the mutation rate of the target region and speed up evolution, while the use of T7 RNA polymerase (T7 RNAP) instead of Cas9-variants broadens the editing window, allowing us to mutate longer regions. Generation of genetic diversity is the foundation of accelerating evolution, and developing mutagenesis tools with different mutation types is of great importance in this process. However, there are still some drawbacks in the above-mentioned techniques, such as low mutation rate and mutation bias, which make them unable to fully meet the needs of users for targeted mutagenesis.

In this work, we developed the targeted *in vivo* mutagenesis system by fusing different DNA-modifying enzymes, cytidine deaminase and T7 RNAP. We obtained three fusion proteins, MAG1/EXO1/REV3-PmCDA1-T7 RNAP, which could significantly increase the mutation rate and expand the mutation spectrum. Our mutagenesis tools can complement the above-mentioned platforms and accelerate the evolution of target genes.

Materials and Methods

Strain construction and cultivation

Escherichia coli -Trans1 T1 was used as the cloning and amplification host. For plasmid construction, the heterologous gene sequences were cloned from the previously published plasmids in Addgene. The DNA-modifying enzymes were PCR amplified from the BY4741 genome. The PCR products were connected to the pRS415 vector. Cloning was performed by SE assembly and BM assembly followed by transformation into *E. coli*- Trans1 T1. *E. coli* strains were cultivated at 37 in liquid LB media (LB, 1% (w/v) tryptone, 0.5% (w/v) yeast extract, 1% (w/v) NaCl) with appropriate antibiotics and selected on LB agar plates with 100μg/mL ampicillin.

BY4741 was used as the chassis strain in this work. The plasmids carrying the mutagenesis cassettes were transformed into yeasts through yeast transformation. The T7-*CAN1* cassette was PCR amplified and integrated into the genome using HIS3 as the selectable marker.Strains were grown at 30 in liquid YPD media (1% yeast extract, 2% peptone, 2% glucose) and selected on SC agar plates lacking leucine, leucine and uracil, or leucine and histidine.

2.2 Canavanine plates selecting for *CAN1* mutant colonies

The mutation rate of *CAN1* (with P_{T7} targeting sequence) was determined to characterize the efficiency of our mutagenesis system in this work. Cells were grown to saturation for 24 h in liquid SC medium lacking appropriate amino acids depending on the autotrophic markers to maintain plasmids. Cultures were then diluted and adjusted to OD_{600}=1 and added into inducing media. Inducing media contains 0.2% galactose and 1 uM β-estradiol. After incubation at 30 for 8˜24 h, 30˜50 μL of the cultures were plated onto SC-Leu-His-Arg plates with 60mg/L canavanine, and the same volume of culture was gradient diluted and plated onto YPD plates. Colonies on canavanine and YPD plates were counted after 2˜3 days to determine the mutation rate.

Mutation rate analysis

The SC-Leu-His-Arg with 60mg/L canavanine plates were incubated at 30 for 2˜3 days and the colonies were counted. The sample size was based on the number of countable colonies on a single plate (up to 500), and the number of colonies on the YPD plates was used as a control. Statistical analysis was performed using Graphpad Prism.

Mutation diversity analysis

For mutation diversity analysis, the average-sized colonies were randomly selected and the target locus was PCR amplified. The PCR products were analyzed by Sanger sequencing and compared with the reference sequence.

Results

PmCDA1 increased the mutation rate of the target gene

Cytidine deaminase PmCDA1 can catalyze the deamination of cytosines, mutating cytosine (C) to uracil (U), while uracil pairs with adenine (A) in the subsequent DNA repair and replication process to complete the C-T conversion.[28, 29] Uracil glycosylase inhibitor (UGI) is usually used to block the activity of uracil glycosylase (UNG) and inhibit the removal of mismatched uracil, thereby reducing other types of mutations such as C-G and C-A.[30, 31] We hypothesized that in the absence of UGI, the types of base substitutions generated by PmCDA1 might be more diverse.[32] We constructed the mutagenesis plasmids carrying the pGAL-PmCDA1-T7 RNAP expression cassette. Meanwhile, we inserted the T7 promoter sequence upstream of the target *CAN1* gene so that PmCDA1-T7 RNAP could be specifically recruited to the target site defined by the T7 promoter (Fig. 1A). The mutation rate was characterized by the frequency of*CAN1* gene inactivation. We performed the assay on yeast strains with and without the mutagenesis plasmids. After induction with galactose for 8˜24 h, we plated cells on canavanine plates capable of inhibiting *CAN1* $^{+}$ cell growth, and counted colonies on canavanine plates to assess the mutation rates. Compared with the control strain, the mutation frequency of the*CAN1* gene in the strain expressing PmCDA1-T7 RNAP was significantly increased (Fig. 1C), indicating that PmCDA1-T7 RNAP can effectively increase the mutation rate of the target gene in *S. cerevisiae* .

Appropriate extension of the linker length can sometimes expand the targeting scope.[30, 33] Based on this, we further investigated the influence of different linker lengths on the mutation effect of PmCDA1-T7 RNAP. Two linker lengths (32a.a. and 84a.a.) were selected (Fig. 1B), and the mutation rate of *CAN1* gene was determined. We observed that the length of linker between PmCDA1 and T7 RNAP had no significant effect on the mutagenic activity of the fusion protein. After 24 h of induction, both PmCDA1-32a.a.-T7 RNAP and PmCDA1-84a.a.-T7 RNAP could increase the mutation frequency up to $1.0\text{x}10^{-3}$˜$1.2\text{x}10^{-3}$(Fig. 1D). We analyzed the mutations generated by PmCDA1-32a.a.-T7 RNAP and PmCDA1-84a.a.-T7 RNAP by sequencing the P_{T7}*CAN1* locus. The data demonstrated that the mutation types generated by PmCDA1-32 a.a.-T7 RNAP and PmCDA1-84 a.a.-T7 RNAP were basically the same, with C-T mutations accounting for more than 97% and the remaining 3% being other types of mutations (Fig. 1E), which was also consistent with the mutation characteristics of PmCDA1.[16] The distribution of the mutations in the *CAN1* gene was also similar (Fig.1F).Therefore, the length of linker between PmCDA1 and T7 RNAP has no significant

effect on the mutation effect. Meanwhile, even without UGI, the mutation types generated by PmCDA1 were really simple, and most of them were C-T mutations.

DNA-modifying enzymes improved mutation effect

When using PmCDA1-T7 RNAP as the mutator, the strong bias towards C-T mutations would reduce the diversity of mutants. In cells, the mismatched U resulting from the deamination of C is excised by DNA-modifying enzymes to form abasic sites. In the subsequent DNA repair process, different bases could be randomly inserted into the abasic sites, resulting in different types of mutations.[34, 35] Thus, we hypothesized that fusing different DNA-modifying enzymes to PmCDA1 would improve the mutation outcome.

We first chose MAG1 as the DNA-modifying enzyme to link to PmCDA1. MAG1 can remove mismatched bases and initiate base excision repair (BER).[36, 37] Overexpression of MAG1 in cells leads to an elevated genomic mutation rate.[38] We assumed that the addition of MAG1 would enhance the excision of mismatched U and create more abasic sites, thereby generating diverse mutation types during the subsequent DNA repair process. Since the relative positions of PmCDA1, T7 RNAP and the DNA-modifying enzymes may influence the mutation outcome, we designed five expression cassettes with different constructions and linkers (Fig. 2A). We compared the mutagenic activity of these fusions with PmCDA1-T7 RNAP and the control strain without mutagenesis fusions. We found that the construction of the fusion proteins significantly affects the mutation outcome. The data indicated that Cons. 3 could raise the mutation frequency up to $1.9x10^{-3}$, which was 1.6 to 2 times higher than that of PmCDA1-T7 RNAP and was the highest among these five fusions. The mutation frequencies of the other four constructions were about $3x10^{-4}$, which was significantly reduced compared with PmCDA1-T7 RNAP (Fig. 2B). When analyzing the types of mutations, we found that C-T mutations accounted for 64.5% of the mutations produced by Cons. 3, followed by G-A mutations (19.2%), C-G mutations (12.5%), and G-C mutations (3.8%). The proportion of non-C-T mutations is 11-fold higher than that of PmCDA1-T7 RNAP (Fig. 2C). We suspect that the increase in non-C-T mutations may be due to the enhancement of the excision of mismatched bases, thus forming more abasic sites-which are important for BER-and increasing the diversity of mutations. Although the mutation frequencies of the other four candidates were low, the mutation types were diverse and most of them were non-C-T mutations. We speculated that the presence of DNA-modifying enzymes in these constructions affected the activity of PmCDA1, resulting in the mutation effect that was apparently different from that of PmCDA1. Considering the mutation frequency and diversity, we selected the Cons. 3 for further work.

Based on this, we chose 6 other DNA-modifying enzymes and analyzed their mutagenic activity (Fig. 3A).[39-44] Among these candidates, EXO1-PmCDA1-T7 RNAP produced the highest mutation frequency of $2.2x10^{-3}$, which was twice that of PmCDA1-T7 RNAP (Fig. 3B). When analyzing the mutations generated by EXO1-PmCDA1-T7 RNAP, we found a strong bias towards C-T mutations, similar to PmCDA1. In 48 randomly selected colonies, C-T mutations accounted for 80.0%, followed by C-G mutations (11.0%) and other types of mutations (8.0%) (Fig. 3C). EXO1 is a key enzyme in DNA double-strand break repair, mismatch repair, and other repair pathways,[43, 45] and we speculated that EXO1 may act synergistically with PmCDA1 to further increase the mutation frequency. In laboratory evolution, a high mutation rate can greatly accelerate the evolution process. The mutation frequency generated by REV3-PmCDA1-T7 RNAP was about $1.26x10^{-3}$, which was slightly lower than that of MAG1-PmCDA1-T7 RNAP and EXO1-PmCDA1-T7 RNAP (Fig. 3B), but the mutation types were diverse, of which C-T mutations accounted for 70.4%, followed by C-G mutations (14.1%), G-A mutations (9.86%) and other mutations (5.64%)(Fig. 3C). REV3 involves in DNA translesion synthesis repair, double-strand break repair, and DNA damage-induced mutagenesis.[44] Therefore, we hypothesized that, similar to MAG1, REV3 strengthens the DNA translesion synthesis repair, in which different bases are inserted into abasic sites, resulting in multiple types of mutations. In the process of laboratory evolution, the occurrence of different types of mutations enlarges the mutant spectrum, and allows us to screen a wider range of desired strains. Different from cytidine-bearing mutators,[21, 22, 24] after altering the mutation spectrum by DNA-modifying enzymes, mutations could occur across all four nucleotides, with G-A or C-G mutations being the main mutation types, except for C-T mutations, meaning that our system is able to play a complementary role to the cytidine-based evolutionary

tools.

Dual T7 promoters increased mutation frequency

After changing the constrution of the fusion proteins and adding DNA-modifying enzymes to improve the mutation effect, the mutagenic activity of mutators have been improved significantly compared with PmCDA1-T7 RNAP. Based on this, we inserted two reverse T7 promoters on both sides of the *CAN1* gene and analyzed the mutation effect of MAG1/EXO1/REV3-PmCDA1-T7 RNAP under this condition (Fig. 4A).[22, 24] We observed that the addition of the second T7 promoter significantly increased the mutation frequency. In the dual T7 promoter system, the mutation rate generated by EXO1-PmCDA1-T7 RNAP could reach 5.13x10^{-3} after 24 h of induction, which was 1.57-fold higher than that of the single T7 promoter system (Fig. 4B). With dual T7 promoters, the mutation frequencies generated by MAG1-PmCDA1-T7 RNAP and REV3-PmCDA1-T7 RNAP were also significantly increased. After induction for 24 h, the mutation rates of strains expressing MAG1-PmCDA1-T7 RNAP and REV3-PmCDA1-T7 RNAP were 3.72x10^{-3} and 3.26x10^{-3}, respectively, which were 2˜2.5-fold higher than that of the single T7 promoter system (Fig. 4B). We speculated that the dual T7 promoters may increase the probability of T7 RNAP binding to the T7 promoter, so that DNA-modifying enzymes-PmCDA1 have a greater chance of acting on the target gene, leading to higher mutation rates.

When analyzing the mutations produced in the dual T7 promoter system, we found that the introduction of the reverse T7 promoter had no strong effect on the mutation types. In the dual T7 promoter system, C-T mutations generated by MAG1-PmCDA1-T7 RNAP comprised 65.7%, followed by G-A mutations (13.6%), C-G mutations (11.9%), and other types of mutations (8.8%). Although the proportion of G-A mutations was slightly lower than that in the single T7 promoter system, other types of mutations increased, such as some transversion mutations like G-T. The mutation types of EXO1-PmCDA1-T7 RNAP in the dual T7 promoter system were barely changed, among which C-T mutations comprised 82.4%, followed by C-G mutations (8.33%), G-A mutations (5.1%) and other mutations (4.17%). In REV3-PmCDA1-T7 RNAP, C-T mutations comprised 73.6%, followed by C-G mutations (12.2%), G-A mutations (12.12%) and other mutations (2.8%) (Fig. 4C). It can be seen that the dual T7 promoter system had little effect on the mutation characteristics of our mutagenesis tools, but slightly increased the frequency of some transversion mutations (such as G-T, G-C, etc.) and made the mutation types more diverse. Existing deaminase-based evolutionary techniques are difficult to achieve transversion mutations, and most of them are biased towards generating specific types of mutations.[22, 24, 25] Therefore, our mutagenesis tools with dual T7 promoters can further enlarge the mutation libraries, thus promoting the process of evolution.

3.4 Εολυτιον οφ κεψ ενζψμες ιν τηε β-ςαροτενε μεταβολις πατηωαψ υσινγ μυταγενεσις τοολς

To explore the potential of our mutagenesis tools in different evolutionary scenarios, such as enhancing the production of valuable compounds, we utilized them in the evolution of the *CrtE* ,*CrtI* and *CrtYB* genes in the β-carotene biosynthetic pathway. We selected the β-carotene-producing yeast strain as the chassis strain and added reverse T7 promoter sequences to *CrtE* ,*CrtI* , and *CrtYB* gene expression cassettes (Fig. 5). After the introduction of mutagenesis plasmids into the chassis strain and induction, we found that there were a few colonies of the strains exhibited different colors. We selected four colonies with distinct color changes and sequenced the target loci. The data suggested that diverse mutations occurred in the *CrtE* , *CrtI* , and*CrtYB* expression cassettes, including transition and transversion mutations. We analyzed these base conversions and found that most of them were C and G mutations, which was also consistent with the mutation characteristics of our mutagenesis tools mentioned above.

We then respectively introduced these single point mutations into the original β-carotene-producing strain and verified whether these strains could still show different colors. We observed that even without the mutagenic fusion proteins, the strains with point mutations still exhibited obvious color changes, indicating that these mutations can effectively alter colony colors and β-carotene yields.

We proved that our mutagenesis tools can generate random mutations in the key enzymes in the β-carotene

biosynthetic pathway, resulting in increased β-carotene yields. These results demonstrated that our mutagenesis tools can be applied to the evolution of non-growth-limiting genes. Even in the absence of the growth pressure or selection, the mutagenesis fusion proteins were able to function robustly.

Discussion

The occurrence of random mutations can improve genetic diversity and play an important role in many fields.[1-4, 6, 9]To date, researchers have developed a handful of mutagenesis techniques to increase the mutation rate and mutation diversity[6]. These tools have produced achievable effects in various chassis cells such as *E. coli* , *S. cerevisiae* , and mammalian cells. In this work, we developed a targeted*in vivo* mutagenesis tool that can significantly improve the mutation rate and broaden the mutation spectrum by fusing DNA-modifying enzymes, cytidine deaminase and T7 RNAP.

We constructed different mutagenesis proteins that could accelerate the evolution of the target gene. We first constructed the plasmid containing the cytidine deaminase PmCDA1-T7 RNAP expression cassette and inserted T7 promoter sequence upstream of the target gene. PmCDA1-T7 RNAP fusions could raise the mutation frequency to about 1.0˜1.2x10^{-3}, and exhibited a strong bias toward C-T mutations, which is consistent with the mutation characteristic of PmCDA1.[21, 22] We hypothesized that enhancing the removal of mismatched U as well as the subsequent DNA repair processes would lead to diverse mutation types.[34, 35, 46] Therefore, we fused different DNA-modifying enzymes with PmCDA1 to improve the mutation effect. The results indicated that the introduction of DNA-modifying enzymes could indeed improve the mutation effect. Among these candidates, MAG1 and REV3 could significantly increase the diversity of mutations. In the mutations generated by MAG1-PmCDA1-T7 RNAP, C-T mutations accounted for 64.5%, followed by G-A mutations (19.2%), C-G mutations (12.5%), and G-C mutations (3.8%). The ratio of non-C-T mutations was 36.5%, which was 11-fold higher than that of PmCDA1-T7 RNAP. REV3-PmCDA1-T7 RNAP could also generate more diverse mutations, of which C-T mutations comprised about 29.6%, nearly 10-fold higher than that of PmCDA1-T7 RNAP. The mutation diversity is crucial to the evolutionary process and the wider mutation spectrum would help us to obtain desired strains more efficiently. Smolke et al. developed TRIDENT system and increased the ratio of non-C-T mutations to about 20%,[23]while Shoulders et al. fused evolved adenosine deaminase to T7 RNAP and developed MutaT7$^{A\text{-}G}$ and eMuataT7$^{A\text{-}G}$, which could generate all transition mutations when being employed with cytidine-bearing mutators.[25] As our mutagenesis fusions could generate higher proportion of C-G and G-A mutations, they could play a complementary role with the above-mentioned tools. Except for the change on the mutation spectra, the addition of DNA-modifying enzymes could also raise the mutation frequency. EXO1-PmCDA1-T7 RNAP could increase the mutation rate up to 2.2x10^{-3}, which is twice as high as PmCDA1-T7 RNAP.

We then added a reverse T7 promoter at the end of the target gene and compared the mutation effects with that of the single T7 promoter system. The introduction of the second T7 promoter could significantly increase the mutation frequency of the target gene. The mutation frequencies of strains with dual T7 promoters were about 1.5˜2.5-fold higher than that in the single T7 promoter system. We hypothesized that the dual T7 promoters might increase the possibility of T7 RNAP binding to the T7 promoter, thus increasing the mutation rate. The proportions of different mutations generated in the dual T7 promoter system were basically the same as the single promoter system. Both Kim et al. and Shoulders et al. have found that the introduction of the second T7 promoter increased the ratios of G-A mutations.[22, 24] However, we did not observe apparent change of the frequency of G-A mutations in the dual T7 promoter system. Instead, we found that the ratios of some transversion mutations (G-T, G-C etc.) increased than that in the single promoter system. We suspect that this difference might be due to the different DNA repair mechanisms in the hosts (*S. cerevisiae* vs. *E. coli*). Our mutators still showed bias toward C-T mutations, and we hypothesized that the introduction of other DNA-modifying enzymes and their combinations might further broaden the mutation spectra.

Finally, we applied our mutagenesis tools to the evolution of the key enzymes in the β-carotene biosynthetic pathway. After induction, we observed apparent color changes compared with the original strain. Most of the mutations were C and G mutations, which was consistent with the mutation characteristic of our

mutators. These results demonstrated that our mutagenesis tools could evolve the non-growth-limiting genes and generate diverse genotypes even without the growth pressure or selection.

Our mutagenesis tools are inducible, so the mutation rate could be flexibly tuned by changing the inducer concentration and induction time. The expression level of our mutagenesis fusion proteins may also influence the mutagenic activity and needs further investigation. In addition, using more DNA-modifying enzymes or their combinations may further alter the mutation effect of the mutagenesis proteins, thus developing evolutionary tools with diverse mutation characteristics to meet different needs of the evolutionary process. Our mutagenesis tools are compatible with continuous evolution, with the help of biosensors or other screening techniques, researchers can efficiently obtain desired mutants. In addition, our mutagenesis tools can also work complementary to the single deaminase-bearing mutators that are biased toward generating transition mutations, effectively creating both transition and transversion mutations, and can be applied to many aspects, such as industrial strain breeding, protein engineering, and so on.

Conclusion

In this work, we established a targeted *in vivo* mutagenesis system by fusing different DNA-modifying enzymes, cytidine deaminase and T7 RNA polymerase. With the introduction of the DNA-modifying enzymes and the dual T7 promoters, our mutagenesis tools could raise the mutation frequency up to 5.13×10^{-3} and significantly expand the mutation spectrum. Besides transition mutations, our system could also efficiently generate transversion mutations, which are difficult to access by cytidine/adenosine-bearing evolutionary tools. Although our system still showed a slight bias toward generating C-T mutations, we hypothesized that this problem would be solved by employing more DNA-modifying enzymes and their combinations. Our mutagenesis tools are effective, flexible, and compatible with multiple evolutionary scenarios. Even without the selection pressure, our system functioned robustly and generated diverse mutations. In conclusion, our mutagenesis system could significantly increase the mutation frequency of target genes and generate mutations of different types, including transversion mutations, providing a powerful tool to accelerate the evolutionary process.

Conflict of interest

The authors declare that there is no conflict of interest.

Acknowledgments

This work was supported by the National Key Research and Development Program of China (2018YFA0900100) and Tianjin Fund for Distinguished Young Scholars (19JCJQJC63300)

Reference

[1] Bracher, J. M., de Hulster, E., Koster, C. C., van den Broek, M., Daran, J. G., van Maris, A. J. A., & Pronk, J. T. (2017). Laboratory Evolution of a Biotin-Requiring Saccharomyces cerevisiae Strain for Full Biotin Prototrophy and Identification of Causal Mutations. *Appl Environ Microbiol* , *83,* AEM.00892-17.

[2] Wang, M., Li, S., & Zhao, H. (2016). Design and engineering of intracellular-metabolite-sensing/regulation gene circuits in Saccharomyces cerevisiae. *Biotechnol Bioeng* , *113* , 206-215.

[3] Reyes, L. H., Gomez, J. M., & Kao, K. C. (2014). Improving carotenoids production in yeast via adaptive laboratory evolution.*Metab Eng* , *21* , 26-33.

[4] Kim, S. R., Skerker, J. M., Kang, W., Lesmana, A., Wei, N., Arkin, A. P., & Jin, Y. S. (2013). Rational and evolutionary engineering approaches uncover a small set of genetic changes efficient for rapid xylose fermentation in Saccharomyces cerevisiae. *PLoS One* , *8* , e57048.

[5] Patzschke, A., Steiger, M. G., Holz, C., Lang, C., Mattanovich, D., & Sauer, M. (2015). Enhanced glutathione production by evolutionary engineering of Saccharomyces cerevisiae strains. *Biotechnol J* ,*10* , 1719-1726.

[6] Wang, G., Li, Q., Zhang, Z., Yin, X., Wang, B., & Yang, X. (2023). Recent progress in adaptive laboratory evolution of industrial microorganisms. *J Ind Microbiol Biotechnol* , *50* .

[7] Hashimoto, S., Ogura, M., Aritomi, K., Hoshida, H., Nishizawa, Y., & Akada, R. (2005). Isolation of auxotrophic mutants of diploid industrial yeast strains after UV mutagenesis. *Appl Environ Microbiol* , *71* , 312-319.

[8] Patnaik, R. (2008). Engineering complex phenotypes in industrial strains. *Biotechnology Progress* , *24* , 38-47.

[9] Jiang, G., Yang, Z., Wang, Y., Yao, M., Chen, Y., Xiao, W., & Yuan, Y. (2020). Enhanced astaxanthin production in yeast via combined mutagenesis and evolution. *Biochemical Engineering Journal* ,*156,* 107519.

[10] Roy, K. R., Smith, J. D., Vonesch, S. C., Lin, G., Tu, C. S., Lederer, A. R., Chu, A., Suresh, S., Nguyen, M., Horecka, J., Tripathi, A., Burnett, W. T., Morgan, M. A., Schulz, J., Orsley, K. M., Wei, W., Aiyar, R. S., Davis, R. W., Bankaitis, V. A., Haber, J. E., Salit, M. L., St Onge, R. P., & Steinmetz, L. M. (2018). Multiplexed precision genome editing with trackable genomic barcodes in yeast. *Nat Biotechnol* , *36* , 512-520.

[11] Bao, Z., HamediRad, M., Xue, P., Xiao, H., Tasan, I., Chao, R., Liang, J., & Zhao, H. (2018). Genome-scale engineering of Saccharomyces cerevisiae with single-nucleotide precision. *Nat Biotechnol* ,*36* , 505-508.

[12] Froehlich, J. J., Uyar, B., Herzog, M., Theil, K., Glazar, P., Akalin, A., & Rajewsky, N. (2021). Parallel genetics of regulatory sequences using scalable genome editing in vivo. *Cell Rep* ,*35* , 108988.

[13] Neggers, J. E., Jacquemyn, M., Dierckx, T., Kleinstiver, B. P., Thibaut, H. J., & Daelemans, D. (2021). enAsCas12a Enables CRISPR-Directed Evolution to Screen for Functional Drug Resistance Mutations in Sequences Inaccessible to SpCas9. *Mol Ther* ,*29* , 208-224.

[14] Tou, C. J., Schaffer, D. V., & Dueber, J. E. (2020). Targeted Diversification in the S. cerevisiae Genome with CRISPR-Guided DNA Polymerase I. *ACS Synth Biol* , *9* , 1911-1916.

[15] Hess, G. T., Fresard, L., Han, K., Lee, C. H., Li, A., Cimprich, K. A., Montgomery, S. B., & Bassik, M. C. (2016). Directed evolution using dCas9-targeted somatic hypermutation in mammalian cells.*Nat Methods* , *13* , 1036-1042.

[16] Nishida, K., Arazoe, T., Yachie, N., Banno, S., Kakimoto, M., Tabata, M., Mochizuki, M., Miyabe, A., Araki, M., Hara, K. Y., Shimatani, Z., & Kondo, A. (2016). Targeted nucleotide editing using hybrid prokaryotic and vertebrate adaptive immune systems.*Science* , *353* .

[17] Despres, P. C., Dube, A. K., Nielly-Thibault, L., Yachie, N., & Landry, C. R. (2018). Double Selection Enhances the Efficiency of Target-AID and Cas9-Based Genome Editing in Yeast. *G3 (Bethesda)* ,*8* , 3163-3171.

[18] Halperin, S. O., Tou, C. J., Wong, E. B., Modavi, C., Schaffer, D. V., & Dueber, J. E. (2018). CRISPR-guided DNA polymerases enable diversification of all nucleotides in a tunable window. *Nature* ,*560* , 248-252.

[19] Feng, S., Liang, L., Shen, C., Lin, D., Li, J., Lyu, L., Liang, W., Zhong, L. L., Cook, G. M., Doi, Y., Chen, C., & Tian, G. B. (2022). A CRISPR-guided mutagenic DNA polymerase strategy for the detection of antibiotic-resistant mutations in M. tuberculosis. *Mol Ther Nucleic Acids* , *29* , 354-367.

[20] Hao, W., Cui, W., Suo, F., Han, L., Cheng, Z., & Zhou, Z. (2022). Construction and application of an efficient dual-base editing platform for Bacillus subtilis evolution employing programmable base conversion. *Chem Sci* , *13* , 14395-14409.

[21] Chen, H., Liu, S., Padula, S., Lesman, D., Griswold, K., Lin, A., Zhao, T., Marshall, J. L., & Chen, F. (2020). Efficient, continuous mutagenesis in human cells using a pseudo-random DNA editor. *Nat Biotechnol* , *38* , 165-168.

[22] Park, H., & Kim, S. (2021). Gene-specific mutagenesis enables rapid continuous evolution of enzymes in vivo. *Nucleic Acids Res* ,*49* , e32.

[23] Cravens, A., Jamil, O. K., Kong, D., Sockolosky, J. T., & Smolke, C. D. (2021). Polymerase-guided base editing enables in vivo mutagenesis and rapid protein engineering. *Nat Commun* , *12* , 1579.

[24] Moore, C. L., Papa, L. J., & Shoulders, M. D. (2018). A Processive Protein Chimera Introduces Mutations across Defined DNA Regions In Vivo. *J Am Chem Soc* , *140* , 11560-11564.

[25] Mengiste, A. A., Wilson, R. H., Weissman, R. F., Papa Iii, L. J., Hendel, S. J., Moore, C. L., Butty, V. L., & Shoulders, M. D. (2023). Expanded MutaT7 toolkit efficiently and simultaneously accesses all possible transition mutations in bacteria. *Nucleic Acids Res* .

[26] Velazquez, E., Alvarez, B., Fernandez, L. A., & de Lorenzo, V. (2022). Hypermutation of specific genomic loci of Pseudomonas putida for continuous evolution of target genes. *Microb Biotechnol* ,*15* , 2309-2323.

[27] Butt, H., Ramirez, J. L. M., & Mahfouz, M. (2022). Synthetic evolution of herbicide resistance using a T7 RNAP-based random DNA base editor. *Life Sci Alliance* , *5* .

[28] Liu, M., & Schatz, D. G. (2009). Balancing AID and DNA repair during somatic hypermutation. *Trends Immunol* , *30* , 173-181.

[29] Muramatsu, M., Kinoshita, K., Fagarasan, S., Yamada, S., Shinkai, Y., & Honjo, T. (2000). Class switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential RNA editing enzyme. *Cell* , *102* , 553-563.

[30] Komor, A. C., Kim, Y. B., Packer, M. S., Zuris, J. A., & Liu, D. R. (2016). Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature* , *533* , 420-424.

[31] Kurt, I. C., Zhou, R., Iyer, S., Garcia, S. P., Miller, B. R., Langner, L. M., Grunewald, J., & Joung, J. K. (2021). CRISPR C-to-G base editors for inducing targeted DNA transversions in human cells.*Nat Biotechnol* , *39* , 41-46.

[32] Zhao, D., Li, J., Li, S., Xin, X., Hu, M., Price, M. A., Rosser, S. J., Bi, C., & Zhang, X. (2021). Glycosylase base editors enable C-to-A and C-to-G base changes. *Nat Biotechnol* , *39* , 35-40.

[33] Zafra, M. P., Schatoff, E. M., Katti, A., Foronda, M., Breinig, M., Schweitzer, A. Y., Simon, A., Han, T., Goswami, S., Montgomery, E., Thibado, J., Kastenhuber, E. R., Sanchez-Rivera, F. J., Shi, J., Vakoc, C. R., Lowe, S. W., Tschaharganeh, D. F., & Dow, L. E. (2018). Optimized base editors enable efficient editing in cells, organoids and mice. *Nat Biotechnol* , *36* , 888-893.

[34] Wallace, S. S. (2014). Base excision repair: a critical player in many games. *DNA Repair (Amst)* , *19* , 14-26.

[35] Chatterjee, N., & Walker, G. C. (2017). Mechanisms of DNA damage, repair, and mutagenesis. *Environ Mol Mutagen* , *58* , 235-263.

[36] Tkach, J. M., Yimit, A., Lee, A. Y., Riffle, M., Costanzo, M., Jaschob, D., Hendry, J. A., Ou, J., Moffat, J., Boone, C., Davis, T. N., Nislow, C., & Brown, G. W. (2012). Dissecting DNA damage response pathways by analysing protein localization and abundance changes during DNA replication stress. *Nat Cell Biol* , *14* , 966-976.

[37] Admiraal, S. J., Eyler, D. E., Baldwin, M. R., Brines, E. M., Lohans, C. T., Schofield, C. J., & O'Brien, P. J. (2019). Expansion of base excision repair compensates for a lack of DNA repair by oxidative dealkylation in budding yeast. *J Biol Chem* , *294* , 13629-13637.

[38] Finney-Manchester, S. P., & Maheshri, N. (2013). Harnessing mutagenic homologous recombination for targeted mutagenesis in vivo by TaGTEAM. *Nucleic Acids Res* , *41* , e99.

[39] Boiteux, S., & Guillet, M. (2004). Abasic sites in DNA: repair and biological consequences in Saccharomyces cerevisiae. *DNA Repair (Amst)* , *3* , 1-12.

[40] Swartzlander, D. B., Griffiths, L. M., Lee, J., Degtyareva, N. P., Doetsch, P. W., & Corbett, A. H. (2010). Regulation of base excision repair: Ntg1 nuclear and mitochondrial dynamic localization in response to genotoxic stress. *Nucleic Acids Res* , *38* , 3963-3974.

[41] Zan, H., White, C. A., Thomas, L. M., Mai, T., Li, G., Xu, Z., Zhang, J., & Casali, P. (2012). Rev1 recruits ung to switch regions and enhances du glycosylation for immunoglobulin class switch DNA recombination. *Cell Rep* , *2* , 1220-1232.

[42] Clairmont, C. S., & D'Andrea, A. D. (2021). REV7 directs DNA repair pathway choice. *Trends Cell Biol* , *31* , 965-978.

[43] Tran, P. T., Erdeniz, N., Dudley, S., & Liskay, R. M. (2002). Characterization of nuclease-dependent functions of Exo1p in Saccharomyces cerevisiae. *DNA Repair* , *1* , 895-912.

[44] Sakamoto, A. N., Stone, J. E., Kissling, G. E., McCulloch, S. D., Pavlov, Y. I., & Kunkel, T. A. (2007). Mutator alleles of yeast DNA polymerase zeta. *DNA Repair (Amst)* , *6* , 1829-1838.

[45] Sertic, S., Quadri, R., Lazzaro, F., & Muzi-Falconi, M. (2020). EXO1: A tightly regulated nuclease. *DNA Repair (Amst)* ,*93* , 102929.

[46] Memisoglu, A., & Samson, L. (2000). Base excision repair in yeast and mammals. *Mutation Research-Fundamental and Molecular Mechanisms of Mutagenesis* , *451* , 39-51.

Figure legends

Fig. 1 PmCDA1-T7 RNAP fusion increased the mutation frequency of the target gene. (**A)** Schematic view of the design of the yeast mutagenesis tool. PmCDA1-T7 RNAP could be recruited to the target gene by the T7 promoter. T7 RNAP would move along the target gene and PmCDA1 could generate random mutations on the target gene. The target gene in this work is the *CAN1* gene, and the T7 promoter is upstream of *CAN1* . (**B)** Construction of PmCDA1-T7 RNAP with different linker lengths and the mutation effects (**C**) after 24 h of induction. **(D)** The mutation rates at *CAN1*in strains expressing PmCDA1-T7 RNAP with different linkers and without PmCDA1-T7 RNAP (null). **(E)** The proportions of different types of mutations at *CAN1* . **(F)** Distribution of mutations at*CAN1* generated by PmCDA1-T7 RNAP fusions with different linker lengths. Values represent the mean and standard deviation of three biologically independent replicates.

Fig. 2 The introduction of DNA-modifying enzymes improved the mutation effect. **(A)** Fusion proteins were designed in five constructions, where Cons. represents the construction. **(B)** The mutation rates at *CAN1* in yeasts with different mutagenesis fusions. Null denotes strains without mutagenesis fusions. (**C**)Fraction of different base substitutions occurring in strains with different mutagenesis fusions. Values represent the mean and standard deviation of three biologically independent replicates.

Fig. 3 Screening of DNA-modifying enzymes to tune the mutation spectra. **(A)** Six other candidates were selected for their involvement in the DNA repair process. DNA-modifying enzymes could significantly alter the mutation rate (**B**) and base conversion types (**C**).

Fig. 4 Dual T7 promoters increased the mutation frequencies.**(A)** The structure of the dual T7 promoter system. The two T7 promoters flanking the *CAN1* gene were in the reverse direction.**(B)** The mutation rates in the dual T7 promoter system were obviously higher than in the single promoter system after 24 h of induction. **(C)** The addition of the second T7 promoter barely changed the proportions of different mutation types. MAG1, EXO1 and REV3 represent strains expressing these fusions with single T7 promoter, while MAG1*, EXO1* and REV3* represent strains with dual T7 promoters. Values represent the mean and standard deviation of three biologically independent replicates. (Student' s t-test, **P<0.01).

Fig. 5 Application of mutagenesis tools to the evolution of key enzymes in the β-carotene biosynthetic pathway. The transcription units of the essential genes in the β-carotene biosynthetic process were flanked by dual T7 promoters. After introducing mutagenesis plasmids and induction, some of the colonies showed apparent color changes. TU denotes transcription unit, including promoter, coding region and terminator. Original strain denotes the β-carotene-producing strain containing T7 promoters and mutagenesis proteins, but without induction.

Fig.1

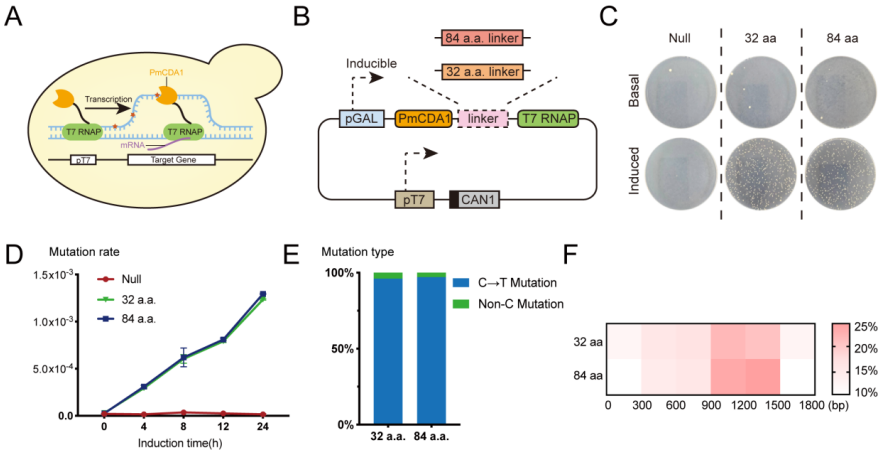

Fig. 2

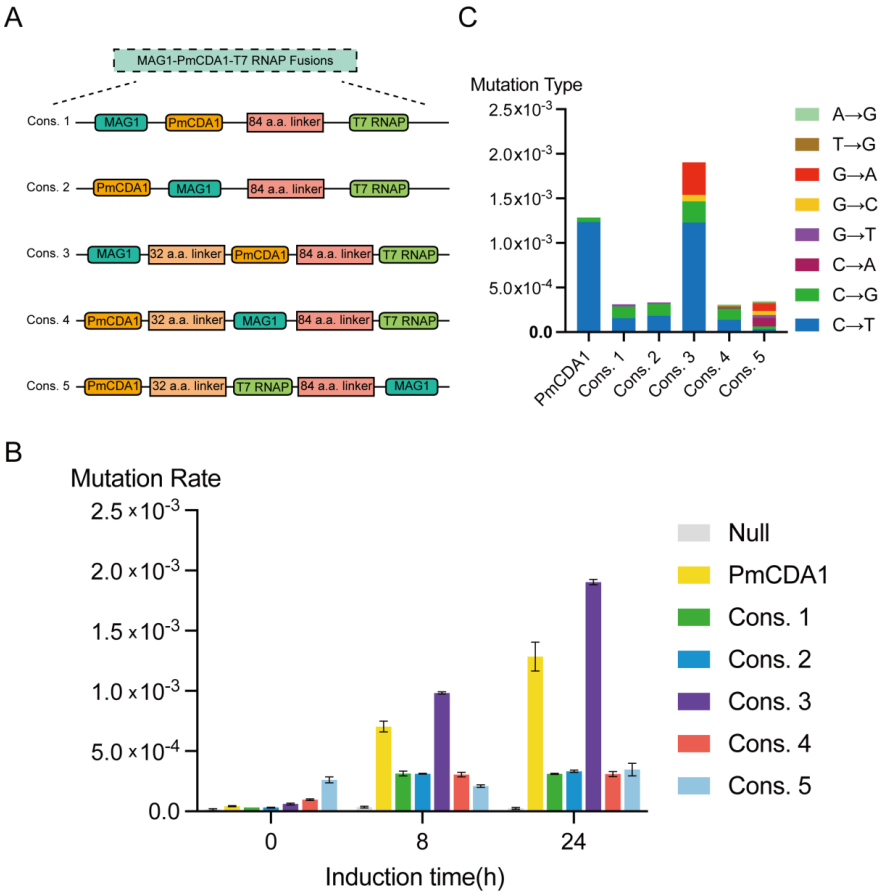

Fig. 3

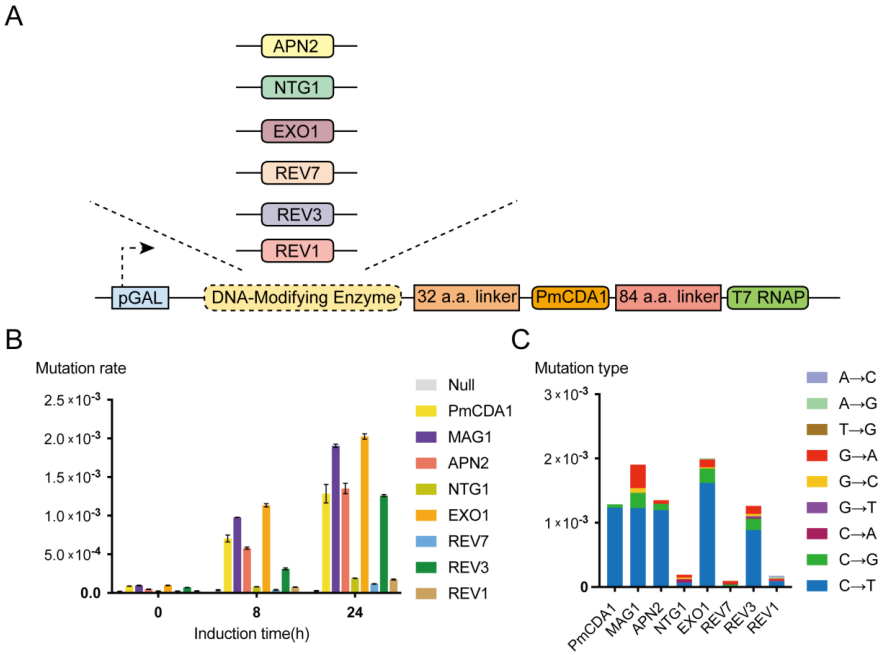

Fig. 4

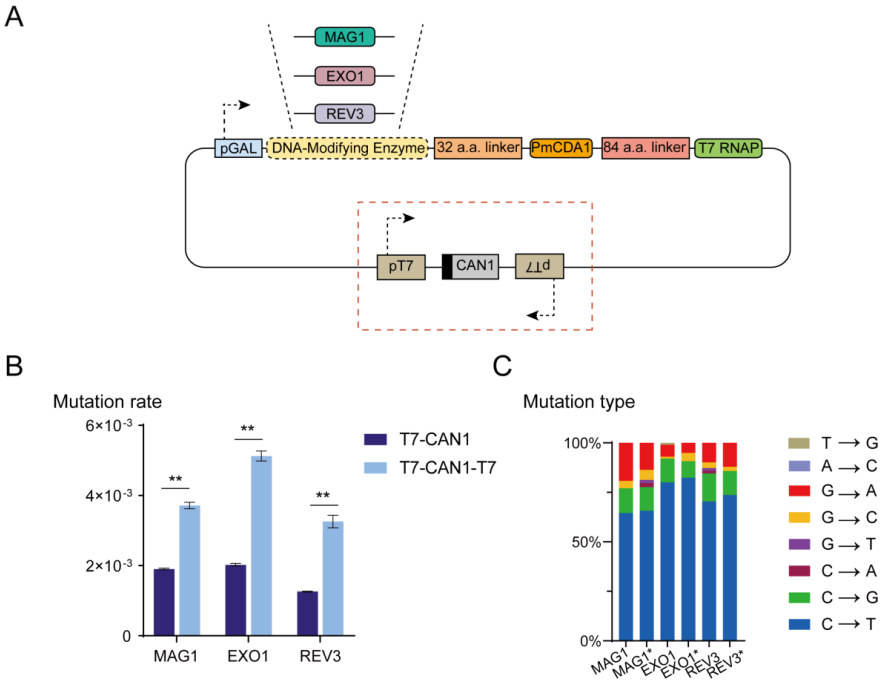

Fig. 5