



# CRISPR-guided DNA polymerases enable directed evolution of luciferase

Weizhe Chen, Rongxiao Zhang, Jiushi Wei, Xuyao Song, Jiali Guo, Xinqi Zou ,  
Zixiang Zhong, Xianghong Wang

## Abstract

**EvolvR, a system that can continuously diversify all nucleotides within a tunable window length at user-defined site, is an attractive method for continuously directed evolution. Our project attempts to conduct directed evolution of luciferase lux of Marine *Vibrio FG-1* isolated from the Yellow Sea by using EvolvR system, aiming to explore a luciferase mutant with better potential performance and various luminescence spectrum. In addition, multiple gRNAs-guided EvolvR system is of great application value. However, due to the problem of genetic instability in the construction process of multiple gRNA, we propose the idea of using ELSA (Extra-long sgRNA Arrays) to help buildmultiple gRNAs-guided EvolvR system. We envision that a stable multiple gRNAs-guided EvolvR system will provide a new strategy for simultaneous evolution of multiple genes, such as metabolic pathways.**

## Introduction

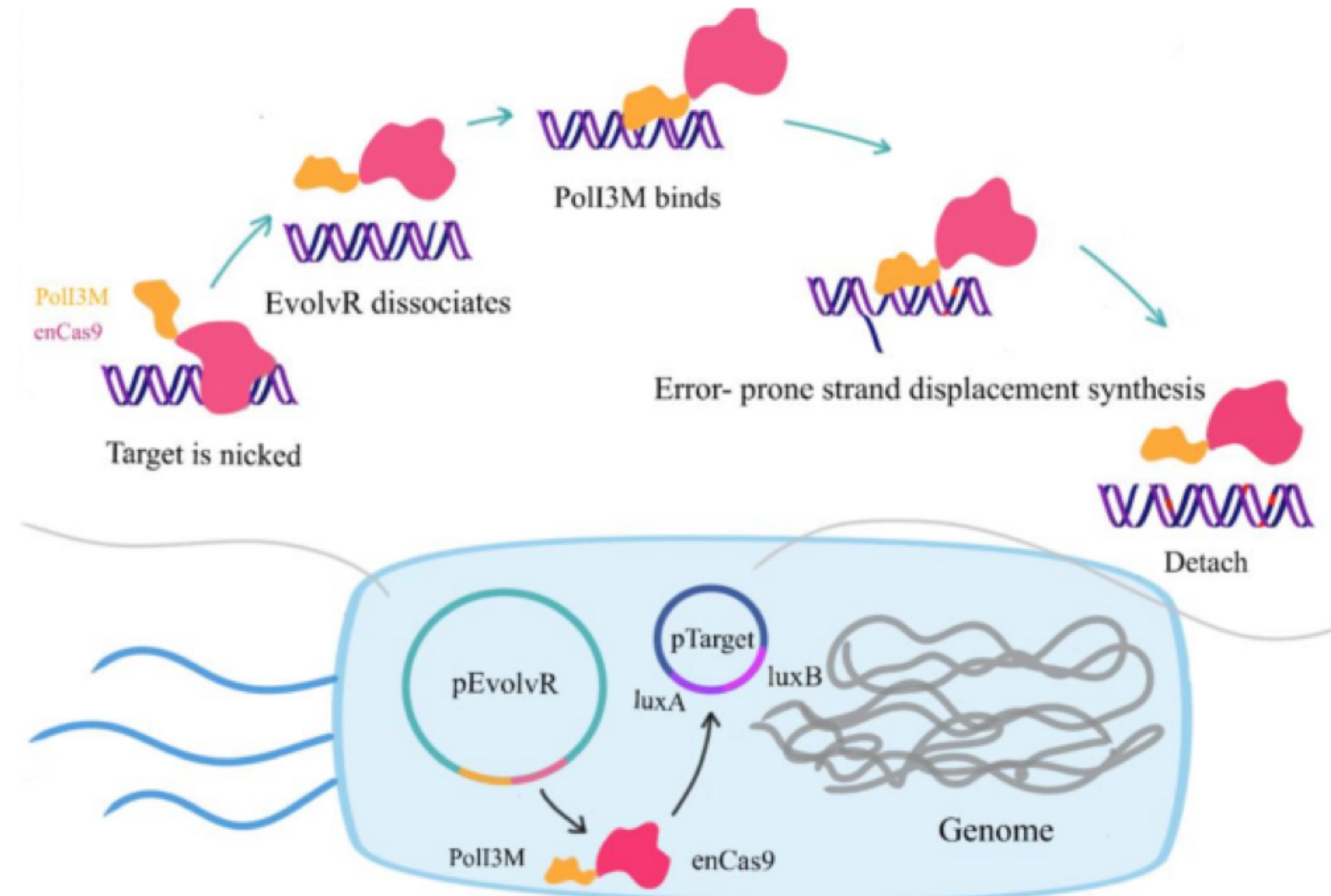
EvolvR system, using low-fidelity DNA polymerases variants as well as CRISPR-guided nickase to mutate the target region, enables the diversification of nucleotides within a tunable window at user-desired loci [1]. It is EvolvR system than can achieve continuous diversification of user defined loci, which will contribute to a wide range of basic and biotechnology applications. After the expression of user-designed gRNA, enCas9 is mediated to target and nick the user-defined loci on the genome or plasmid. The pol13M mutant with low fidelity degrades single

strand from 5' to 3' from the nick and synthesizes a new strand replace with a length of about 350bp (Figure 1a). The EvolvR system enable targeted genes to mutate, express in vivo, screening the mutants, and recycle these steps to achieve continuous-mutagenesis. In comparison to other directed evolution techniques for bacteria (MAGE, MMR, error-prone PCR), EvolvR is a continuous directed evolution platform, which does not require to build mutation libraries and specially evolve user-desired locis with high mutation rates [2,3]. Compared with other directed evolution methods, it can spontaneously mutate target gene in an efficient and tunable way without harming the host bacteria, which is also labor-saving. Recently, the EvolvR system has now been demonstrated to work in eukaryotes such as yeast [4]. This will broaden the range of applications of the system and contribute to a wide range of future biotechnology applications [5].

Bacterial luciferase is a flavin-dependent monooxygenase, catalyzing the reaction of O<sub>2</sub> with reduced FMN and a long-chain aliphatic aldehyde, and releasing oxidized FMN, water, and a fatty acid, with concomitant emission of blue-green light [6]. Luciferase substrate aliphatic aldehyde is with low toxicity and easy to enter cells, so it is harmless to cells in most cases. In comparison with GFP, luciferase does not require external light source, and the process of luminescence detection does not affect the metabolic activities of cells, which is suitable for long-term longitudinal studies in living cells, tissues and animals, especially light-sensitive tissues like retina. Luciferase produce either diffusion or in situ aggregation, which means the expression of luminescence can be employed to accurately locate cells [6].

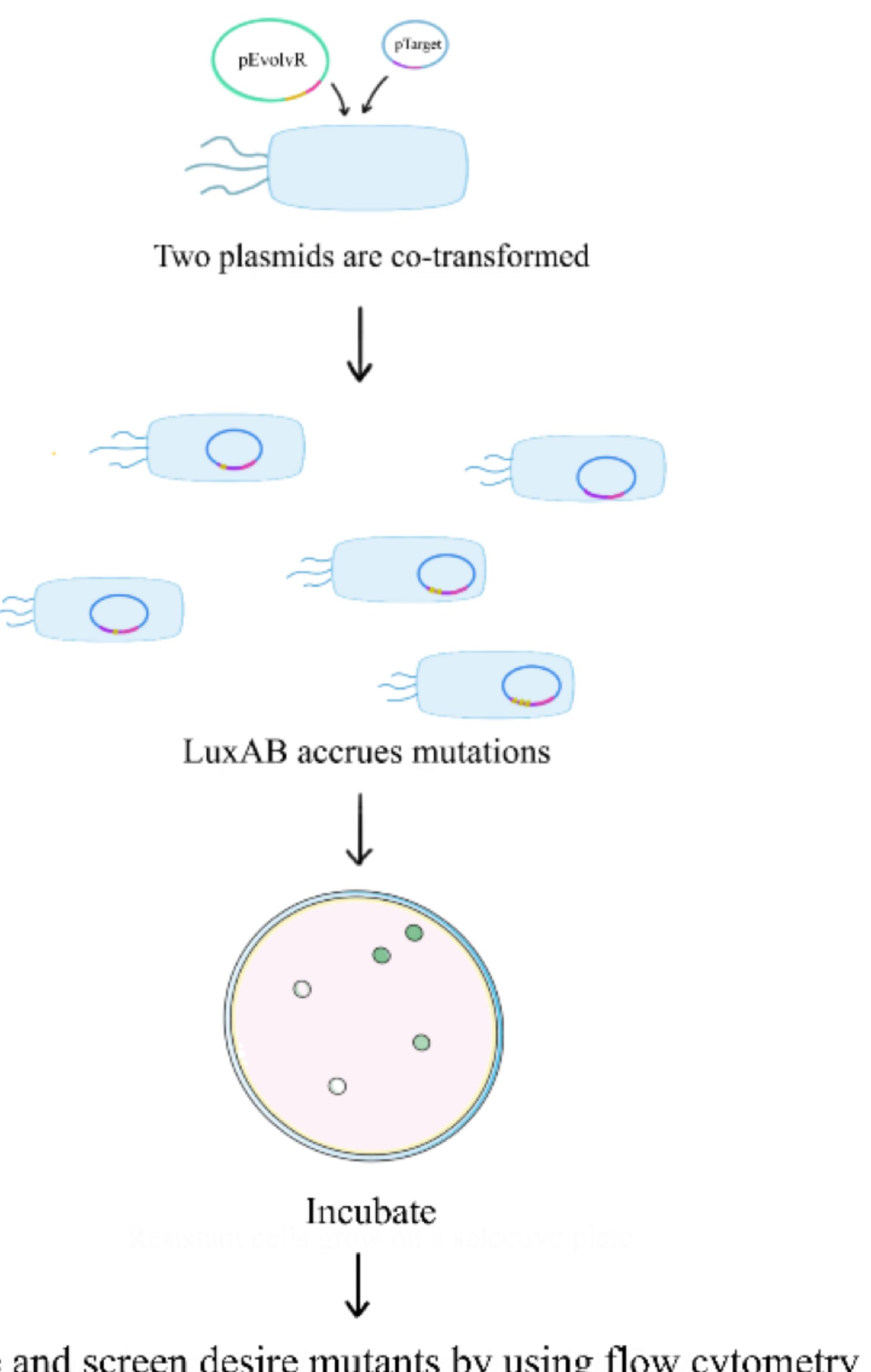
In our study, the *Vibrio FG-1* isolated from the Yellow Sea was identified and its bioluminescence related genes luciferase *luxA* and *luxB* were isolated. The luminescence system was reconstructed in *Escherichia coli*. By designing gRNA targeting different sites of *luxA*, the CRISPR-mediated EvolvR system is constructed to carry out directed evolution of luciferase  $\alpha$  domain, aiming to explore potential luciferase mutants with better performance and various luminescence spectrum, potentially contributing to the field of optogenetics and biological imaging. EvolvR system could couple the EvolvR-mediated mutagenesis to a genetic screen of non-selectable phenotype. Shakked O. Halperin et al. have demonstrated the ability of this system by using GFP protein mutation experience, which reveals that it is feasible to use this system for directed evolution and screening of luciferase mutants[1](figure 1b).

In addition, multiple gRNAs-guided EvolvR system is of great application value. Optimization is worthy of consideration. The EvolvR system with multiple targeting sites has difficulties in the construction of sgRNAs [7]. We propose to use ELSA (Extra-long sgRNA Arrays) to help design and build the EvolvR system with multiple sgRNAs. The system will be able to target multiple loci simultaneously, co-directed evolve series of genes, such as entire metabolic pathways. Also, we put forward an idea that the mutagenesis window length of the EvolvR system can be enlarged by using several sgRNAs targeting adjacent loci.



*Fig-1a.*A brief description of EvolvR

The system includes a low fidelity DNA polymerase I (PolI3M), and a gRNA-mediated nuclease (enCas9). After gRNA binds to a specific site, enCas9 cleave a nick and detach, then PolI3M binds to the nick and resynthesizes downstream strand from 5' to 3'. EvolvR is a two-plasmid system, plasmid pEvolvR involves enCas9-PolI3M-TBD and sgRNA, plasmid pTarget carries the target gene.



*Fig-1b.*Expression and screening of mutants

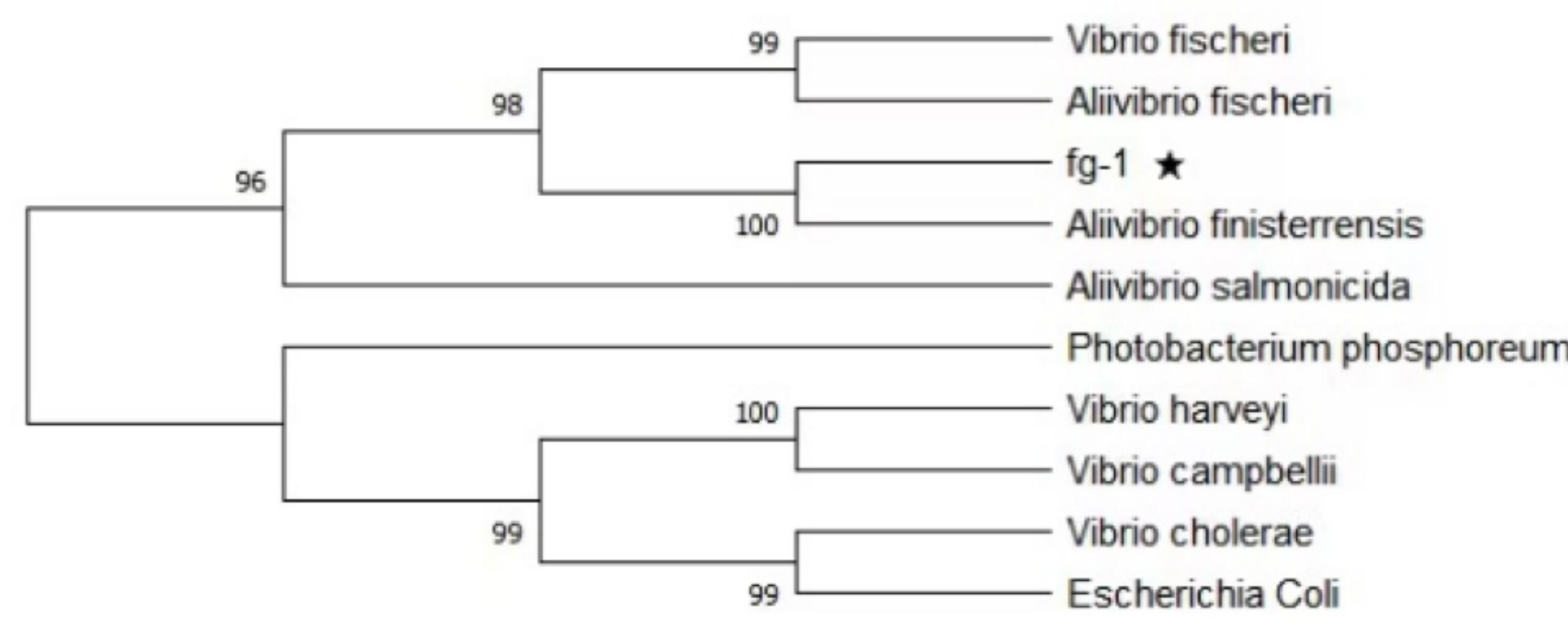
The flow chart of evolving and screening the mutants by EvolvR system. Flow cytometry is used to test the wavelength of luciferase mutants.

## Result

### 1.Species identification and structural

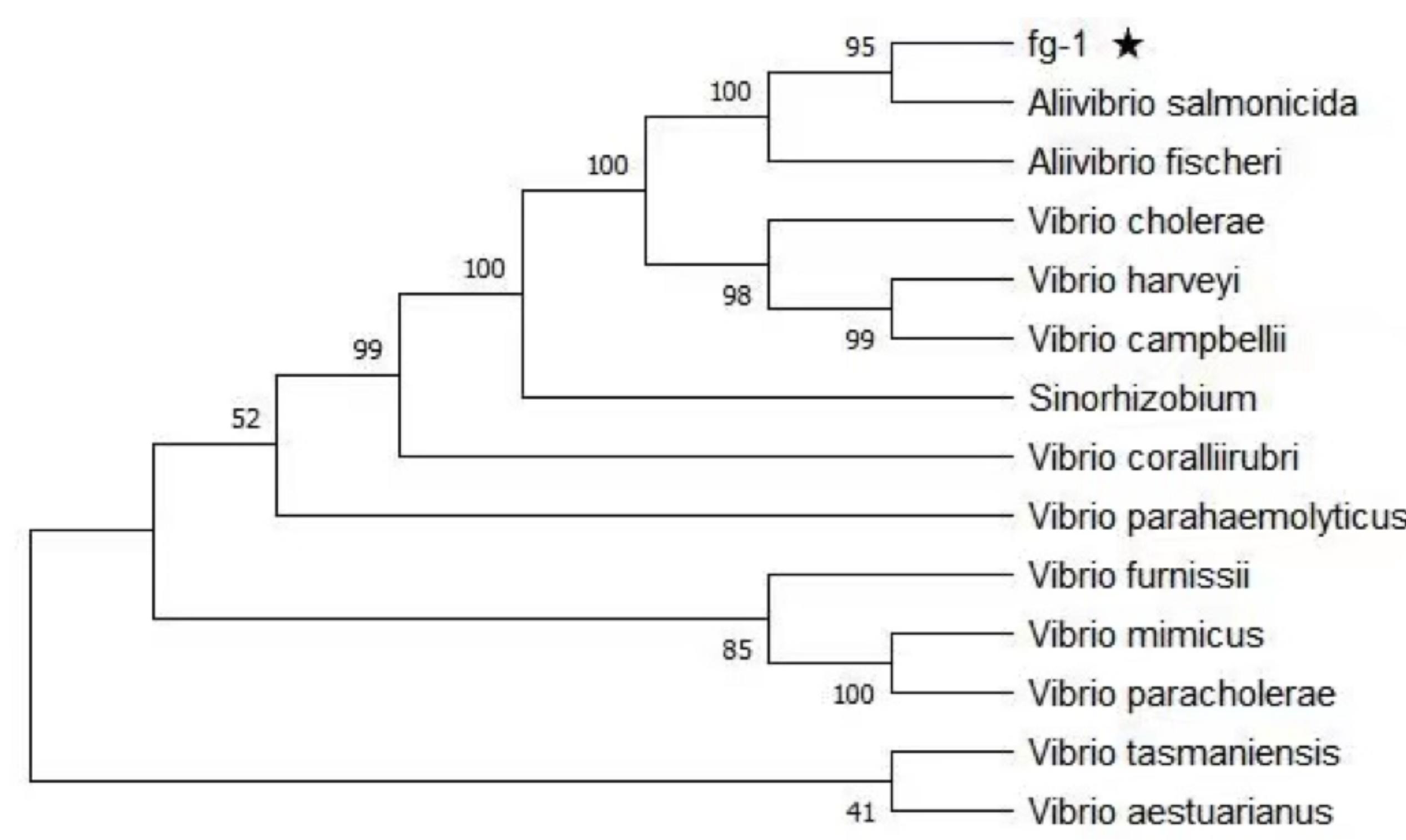
## **prediction**

We constructed phylogenetic tree of 16S rRNA gene to identified a cyan light-emitting luminescent bacteria strain, *Vibrio FG-1*, which was isolated from the Yellow sea. Close phylogenetic relationship between *Allivibrio finisterrensis* and *Vibrio FG-1* was reflected by MEGAX analysis. (Figure 2a) Phylogenetic tree construction of *luxA* has shown that the *luxA* of *Vibrio FG-1* is closely related to *luxA* of *Aliivibrio Salmonicida* and *Aliivibrio Fischeri* (Figure 2b).



*Fig-2a.FG-1's Phylogenetic tree of 16s*

Phylogenetic tree was constructed based on 16S rRNA genes. *Vibrio FG-1* belonged to Allivibrio group with a confidence level of 96% and has a high phylogentic relationship with *Aliivibrio finisterrensis*.

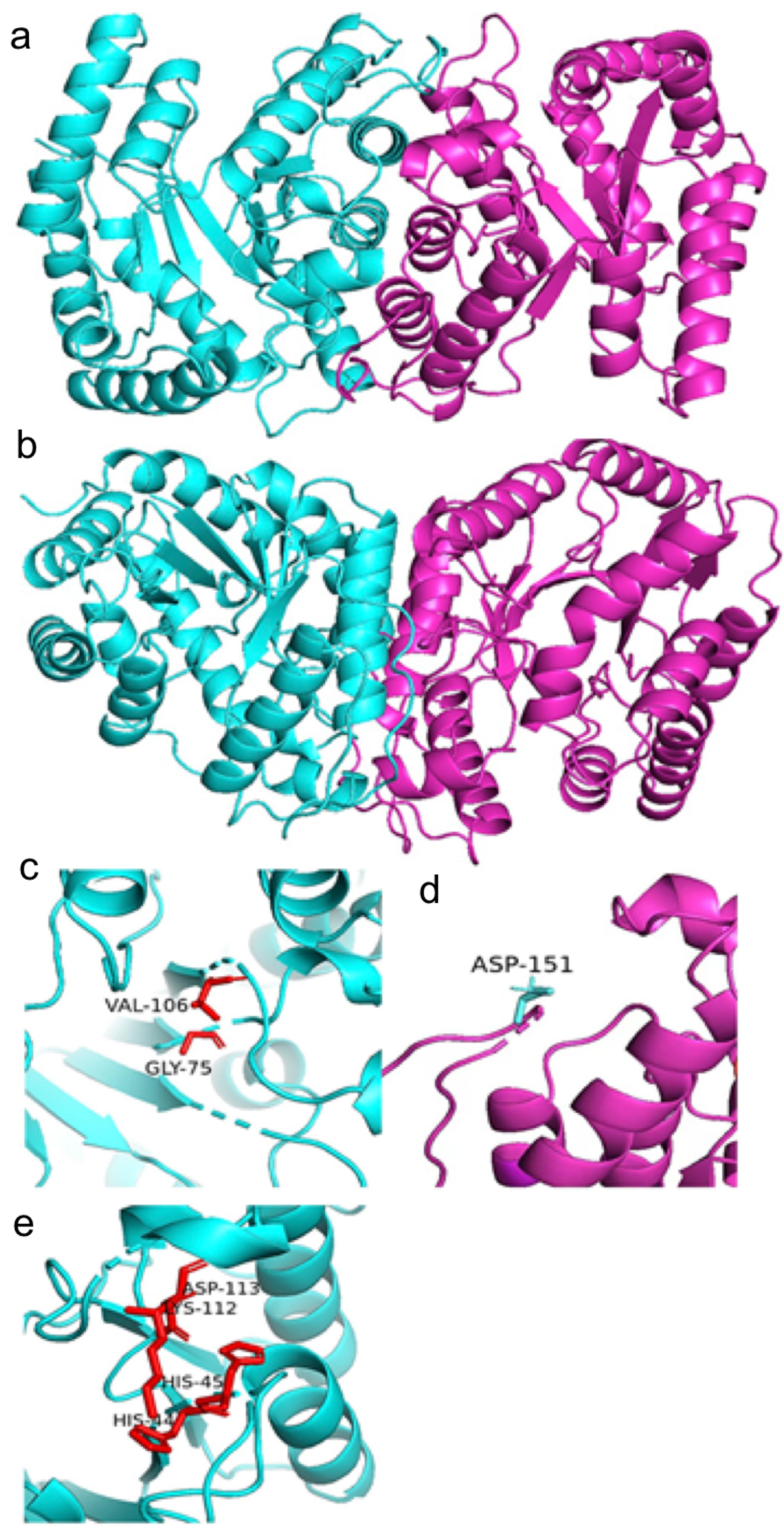


*Fig-2b.lux protein's Phylogenetic tree of FG-1*

Phylogenetic tree construction of *luxA* has shown that *Vibrio FG-1* is closely related to *Aliivibrio Salmonicida* and *Aliivibrio Fischeri*, with a confidence level of 95%.

We designed primers based on luciferase Lux of sequentially similar species, and amplified *luxA* and *luxB* genes of luciferase. In the predicted structure of the  $\alpha$  and  $\beta$  subunit of the luciferase we obtained, similarities to the same subunits of *Vibrio harveyi* were observed. Under this circumstance, we hypothesized that this two subunits will combine with each other and form an asymmetrical structure (Figure 3a, b). The Asp113, RHis44, RLys112 and RHis45 at the  $\alpha$  subunit we predicted

(Figure 3e), which will electro statistically interact with each other after binding to FMN [8], indicates that this  $\alpha$  subunit will also bind to FMN. In this luciferase, the FMN binding site at  $\alpha$  subunit, Ala75, mutated to Gly (Figure 3c), but considering that this situation also exists in some bacterial luciferase [9], we could not confirm the influence of this substitute of amino acid. In addition, mutations in Cys106 at  $\alpha$  subunit may lead to the instability of intermediate 4a,5-dihydro-4a-hydroperoxy flavin [10] and further reduce the enzyme activity. Also, the Substitutions Y151D on the  $\beta$  subunit (Figure 3d) may also lead to reductions in luciferase activity and total quantum yield [8]. Besides, A series of mutations occurred in amino acid residues on Mobile loop -- and F at position 272 was mutated to H, which did not appear in previous studies. Considering that mobile loop is the boundary between the active center and the outside world [10], the mutation of this contact site may lead to the decrease of enzyme activity, which deserves more attention. Furthermore, A series of mutations occurred on the mobile loop of  $\alpha$  subunit, including an F272H, which has not been reported before. Given that the mobile loop worked as a boundary between the active center and the solvent [8], the mutation of this contact site may lead to the decrease of enzyme activity and deserves more attention.



**Fig-3. The structure of the  $\alpha$  and  $\beta$  subunit of the luciferase**

a, b. The asymmetrical dimer formed by this two subunits. c, d .Substitutions of some active site at  $\alpha$  and  $\beta$  subunit that may lead to decrease in enzyme activity. e. The Asp113, RHis44, RLys112 and RHis45 at the  $\alpha$  subunit, which will interact with each other after binding to FMN.

## 2. Construction of EvolvR system

We attempt to express luciferase in *Escherichia.coli* DH5 $\alpha$  and directed mutate luciferase by using enCas9-Pol13M-TBD mediated EvolvR system. We construct and transform pTarget plasmid with luciferase gene and pEvolvR plasmid with sgRNA mediated EvolvR system into *Escherichia.coli* respectively. By observing the phenotype and testing wavelength of luciferase mutants, screening and amplifying, luciferase mutants with desired characteristics will be finally obtained [1].

### 2.1 Construction of pTarget

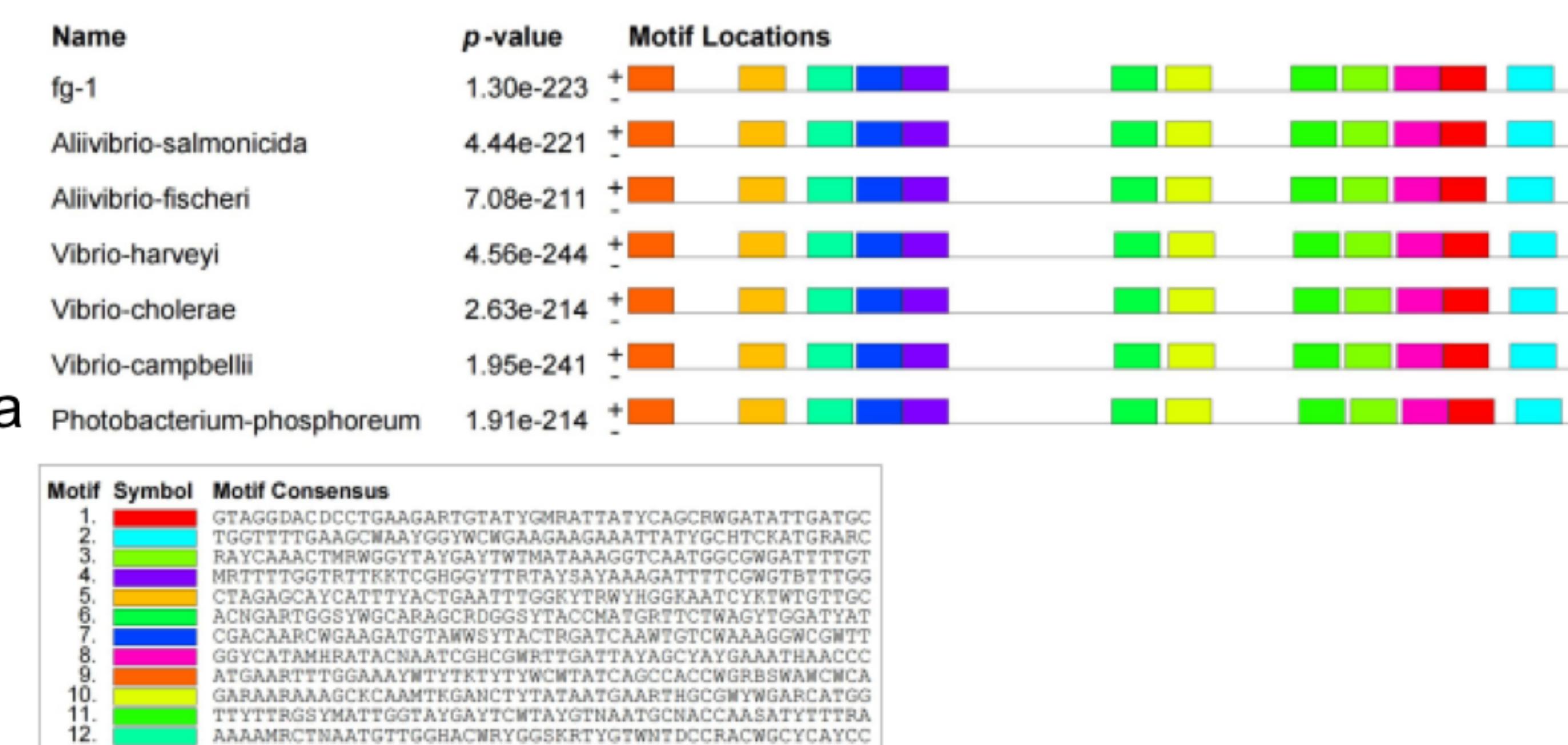
We amplified luxA and luxB genes from *Vibrio FG-1* by PCR, and added J23119 promoter to them. LuxAB fragment, AmpR fragment and P15A ori

fragment were assembled by seamless cloning method. Three fragments with equal molar ratio and Seamless cloning MasterMix was added to incubate according to the instructions of seamless cloning kit. About 20bp homologous sequence between each fragment permit the construction of plasmid. Single colony will be screened on plate containing ampicillin, further colony PCR will be used to verified.

## 2.2 Design of sgRNA

We selected some sgRNA candidates to target luxA gene for directed mutation. sgRNA selection is mainly based on the following points:

- Expression of two gRNAs that nick separate strands at genomic loci separated within 100 bp was lethal, whereas nicking the same strand at this 100-bp distance was not lethal.
- In this project, we used enCas9, a mutant of spCas9 (H840A). 20BP-NGG-spCas9 was used for Protospacer Adjacent Motif (PAM) when the online software CRISPROR was used to select and design gRNA [11, 12].
- Avoiding GC levels that are too high (above 80%) or too low (20%)
- Select gRNA with low off-target rate
- Due to uncertainty of the specific sites that potentially determine the change of luminescence wavelength of luciferase, we analyze the conserved domains of  $\alpha$  unit of luciferase by MEME.



**Fig-4. The conserved domain of the amino acid sequence of luxA protein**

- the comparison of the conserved domain of the amino acid sequence of luxA protein of Vibrio fg-1 with other vibrio luxA. Rectangles with the same color represent the same motif.
- specific amino acid sequences of motif represented by each color rectangle.

According to above several rules, we design the following five gRNA candidates (figure 5a). Due to uncertainty of the specific sites that potentially determine the change of luminescence wavelength of luciferase, we design some gRNA targeting the active sites, and others targeting the conserved region (figure 5b).

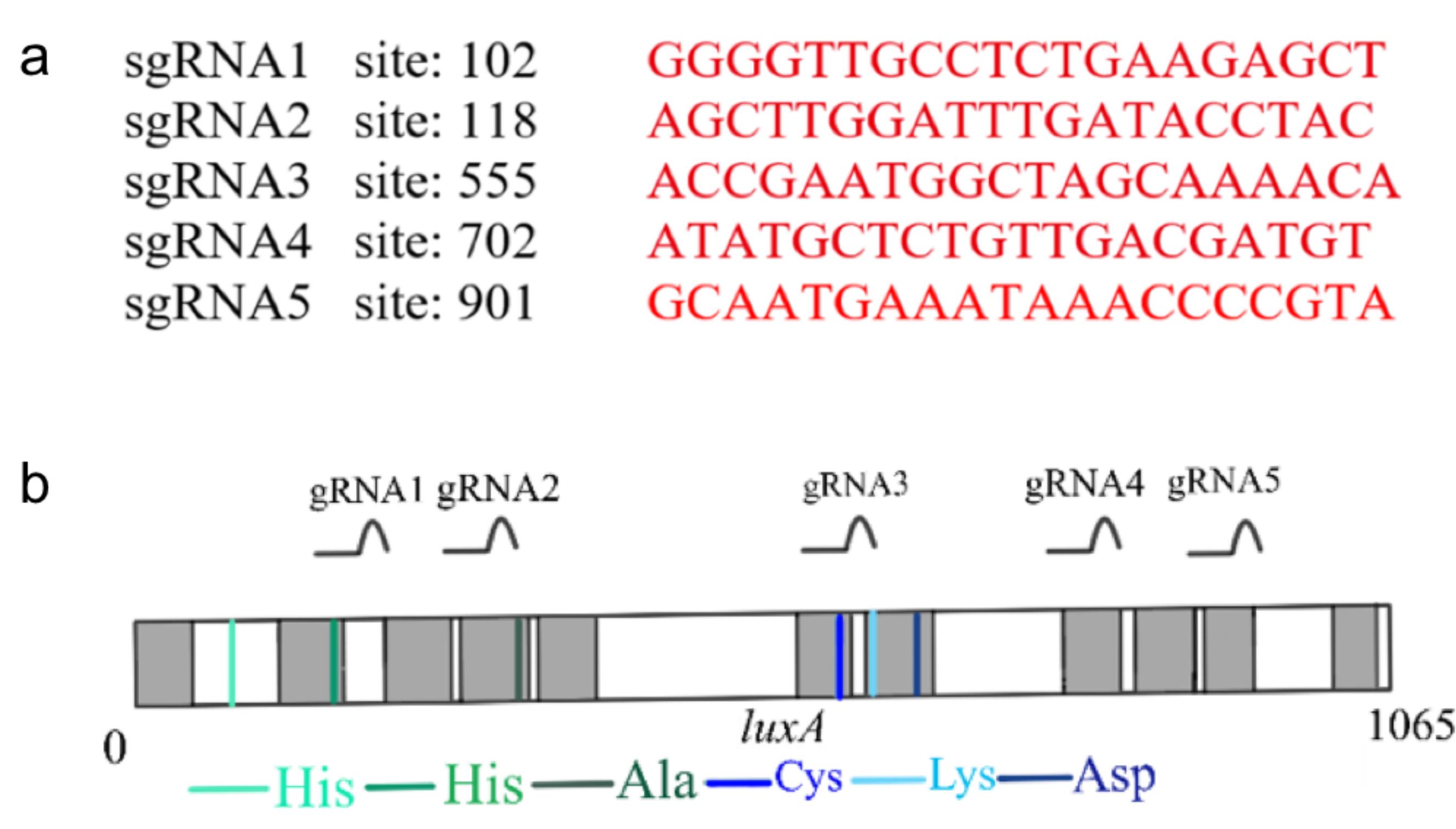


Fig-5 .sgRNACandidates

The gray bands in the figure represent the conserved regions of luxA gene, and the colored lines mark the corresponding active sites of luciferase. We designed five gRNAs targeting luxA gene to mutate luxA individually or simultaneously.

### 2.3 Construction of pEvolvR

We inserted promoter J23119 and gRNA in front of the sgRNA scaffold on the pEvolvR plasmids which were purchased from Addgene website by recombination. Five pEvolvR plasmids were separately constructed for the five gRNA candidates. After E.coli was introduced, the expression of enCas9-Pol13M-TBD was induced by dehydrating tetracycline.

### 3. Fluctuation analysis assay

To sensitively quantify the mutation rate, so as to verify that the mutagenesis is created by EvolvR system instead of higher basic mutation rates, we designed a Fluctuation analysis to verify. In order to optimize the culture scale, different glucose concentrations were applied to control the number of cell divisions when the culture reached saturation. Based on the colony number of gentamicin-sulfate resistant LB plate, the plate has no colony when the dilution ratio is  $10^{-6}$ . At this time,  $OD_{600} = 0.05$  means colony number is  $510^6$  CFU / ml ( $N_0$ ). After diluting to  $10^6$  CFU / 30ml, the

plate is added to the LB liquid medium with different glucose concentration and coated on the complete LB medium. According to the experimental steps in Methods, when glucose concentration was 0.1%, 0.05%, 0.01% and 0.005%, the proportion of unmutated plates and total plates was 44.7%, 30.6%, 89% and 64.4% respectively.  $P_0$  method (Poisson distribution) requires that the percentage of culture medium ( $p_0$ ) without mutants should be between 10% and 80%, so 0.01% LB liquid culture medium with glucose concentration was selected. To calculate the viability and the mutation rate of bacterial according to the formula in Fig-7, we need to calculate the number of bacteria that have mutated (**a**), the total number of bacteria after culture ( $N_f$ ) and the actual number of viable bacteria (**b**) [13].

$$\text{viability} = \frac{b-a}{N_f - a}$$

$$p_0 = \frac{N_f - c}{N_f - a}$$

$$u = \frac{-\ln(p_0)}{\text{vability} \times N_f}$$

Fig-6. The formula of fluctuation test

The colony number (**a**) is 1.3 CFU per plate. After incubation for one day in the 96-well plate,  $OD_{600} = 0.0699$ , meaning total number of bacteria  $N_f = 210^5$  CFU. The actual number of viable bacteria coated in the complete culture medium(**b**) is 322.5 CFU per plate, and the **viability** of bacterial was calculated ( $\text{viability} = 0.0016$ ). The number of mutated cells ( $c = 28.64$  CFU) was obtained when they were coated in selective medium. Therefore, zero-level events -  $p_0 = 0.99986$ . The mutation rate -  $u = 4.3810^{-7}$ . The original mutation rate is high, and the analysis results are inaccurate due to fewer samples and fewer repeated experiments.

### 4. Multi-gRNAs

In industrial production, the desired product or phenotype often requires the expression of multiple genes to be co-regulated, such as changing a metabolic pathway, or try to study complex gene regulatory networks or polygenic diseases. The addition of multiple gRNAs to the EvolvR system will enhance the utility of the system. If EvolvR can target multiple loci of a family of genes and simultaneously direct the evolution of a family of genes, it has the potential to complete the co-directed evolution of an entire metabolic pathway or family of genes in a short period of time. EvolvR can increase the overall length of the mutation window by targeting mutations at different loci of the same gene, and could even customize the length of the mutation window by selecting the locus and number of gRNAs. In previous researches, the EvolvR system has been proved to make simultaneous diversification of genomic loci through co-expression of multiple gRNAs. The EvolvR system allows multiple gRNAs to simultaneously target two genes for targeted mutations and screen for mutants with two target phenotypes. However, studies have indicated that if multiple gRNAs were expressed at the same time, those repeated DNA sequences lead to increased difficulty in plasmid assembly and decreased genetic stability, especially in organisms with frequent homologous recombination [14, 15]. Alexander C et al. designed and screened a series of non-repeating sequence toolbox to help build multiple sgRNA Arrays targeting different loci [7]. The toolkits consist of a sequence non-repeating series of promoters and spacers, and a series of non-repeating gRNA scaffolds designed through RNA structure prediction and Monte Carlo optimization, as well as experimental testing. In our design of the gRNA, we attempt to use nonrepetitive extra-long gRNA arrays (ELSA) from the literature to construct a series of gRNA arrays targeting different loci of luciferase genes (figure 7c) [7].

a

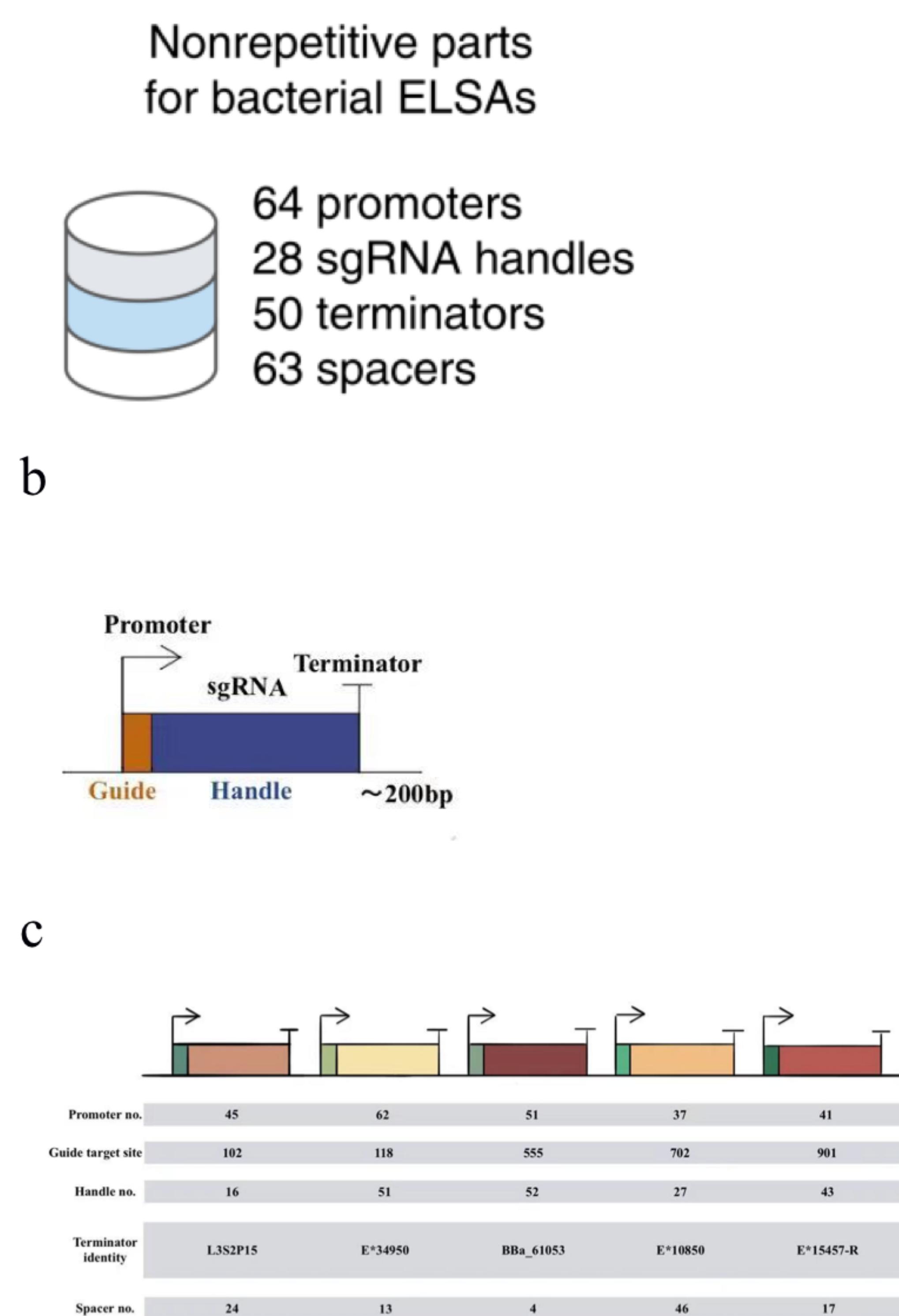


Fig-7.

- a. the non-repetitive extra-long sgRNA arrays literature, including promoters, gRNA handles, spacers and terminators. Different combination of these parts can help avoid repetition.
- b. the circuit of expressing gRNA.
- c. using some of parts candidates in ELSA library to construct the circuit of co-expressing our 5 gRNA candidates.

## Discussion

Current applications of luciferase have focused on its use as a tracer or reporter gene in cytology or genetic expression studies, to mark the movement of cellular structures and gene expression products that are difficult to detect by conventional methods or at higher cost.

However, the current wild-type luciferase, either in terms of luminescence intensity or wavelength variety richness, is inadequate as a tool for more

complex cytological or genetic studies. In our study, the luciferase gene *luxAB* from *Vibrio FG-1* was mutated by the EvolvR system. By precisely selecting the gene sites that may determine the different characteristics of the enzyme, we can mutate luciferase's luminous wavelength and luminous intensity to make it as a reporter gene or a tracer molecule. The higher luminescence

intensity enables researchers to detect and screen targets in a simpler and cheaper way in cytological and genetic studies using luciferase as a reporter gene, or when researchers need to simultaneously detect multiple gene expression products in a biological expression system, luciferase, which can emit fluorescence at different wavelengths, can function as reporter genes of different genes under test.

As a result, the correct object can be screened at one time through spectral analysis, reducing the difficulty and cost of screening. In addition to the above application prospects, luciferase itself substrate and produce by-products for harmlessness cells, and the universality of the protein structure and sequence, the gene and even can not only in prokaryotic system or simple eukaryotic systems (such as yeast) in the expression of successful and efficient, even in the more advanced mammalian cells (HeLa), also can be used as a highly efficient report gene [16]. If the wavelength of luciferase is increased by directed evolution, it will be helpful for it to be better observed in animal tissues. Moreover, luminescence depends on metabolic energy and could only be seen in metabolically active cells, thus preventing artifacts caused by the observation of severely damaged or dead cells. In addition, for the improvement of luciferase Lux, we also proposed the use of a BRET system (bioluminescence resonance energy transfer system) to optimize luciferase. Luciferase was used as light source and fluorescent protein becomes energy receptor to enable greater adjustability of wavelength and intensity of luminous systems.

Meanwhile, we have attempted to use the EvolvR system to carry out multi-locus, precisely targeted directed evolutionary mutation of *lux* gene. As the EvolvR system for single gRNA guided by the single site targeted mutation possesses high efficiency according to the theory, we attempted to extend the length of the mutation window by connecting five highly expressed gRNAs designed to simultaneously target different target sites of the same gene for random mutations, providing

theoretical guidance for the rapid construction of large mutation libraries using EvolvR system. In future experiments, we plan to further verify the validity of this theory through experimental methods, which would strongly prove the application value of the EvolvR system in the field of directed evolution of non-antibiotic genes such as enzyme genes and new drug development.

In a word, the EvolvR system helps to evolve the structure and function of proteins, while the directed evolution of luciferase will greatly enrich the biospectrum, making it a more widely used reporter gene. Future work will enable luciferase genes to be expressed more efficiently in different types of cells through directed evolution, making it easier to apply in cytology, genetics and other disciplines.

## Materials and Methods

### 1. Seamless cloning MasterMix constructs plasmid

#### 1.1. DNA manipulation.

PCR was performed with 2×Taq PCR MasterMix (Solarbio, Beijing, China). Plasmid DNA was isolated with the TIANprep Mini Plasmid Kit (TIANGEN, Beijing, China). DNA fragments were purified from agarose gels by using the Universal DNA Purification Kit (TIANGEN, Beijing, China) or the Monarch DNA Gel Extraction Kit (NEB). DNA sequencing and primer synthesis were carried out by Sangon Biotech (Shanghai, China).

#### 1.2 Plasmid construction.

- a. The plasmids and primers used in this study are listed in the supplemental material.
- b. To construct pTarget-luxAB, the coding regions of luxAB, p15A, and AmpR are amplified from the genome of Fg-1, a plasmid from 2021OUC-China iGEM and Cloned UpB\_4A3m with three different primer pairs luxAB-F/luxAB-R,

- p15A-R/p15A-F, AmpR-R/AmpR-F, respectively. The J23119 promoter is added to PCR products of luxAB by primersluxAB-F and J23119R, resulting in J&luxAB. Oligonucleotides are designed to contribute flanking homologous regions to adjacent DNA fragments of 20-25bp in length, resulting in p15A-homo and AmpR-Homo.
- c. A new pEvolvR plasmid expressing enCas9-Poli3M-TBD with gRNA targeting luxAB is constructed. Three fragments pEA, pEB and pEC with flanking homologous regions to adjacent DNA fragments of 20-25bp are amplified from pEvolvR-enCas9-Poli3M-TBD with primer pairs pEA-R/pEA-F, pEB-R/pEB-F, pEC-R/pEC-F, respectively. Double-stranded gRNA is gained after DNA annealing, and it is modified by the promoter with the primer pair—. Similarly, then homologous regions are added to produce J&gRNA-Homo.
  - d. Two plasmids were both constructed using the Seamless cloning Master Mix provided by Sangon Biotech (Shanghai, China). The plasmids pTarget-luxAB and pEvolvR were cotransformed into E. coli strainDH5 $\alpha$ , following the specification steps provided.

## 2. Structural forecasting

The nucleotide sequences of gene luxA and luxB of fg-1 were translated into amino acid sequences from the first promoter. The structures of both luxA and luxB protein were predicted by phyre2 (<http://www.sbg.bio.ic.ac.uk/phyre2/html/page.cgi?id=index>) through homology modeling method. Futher analysis of protein structure including the alignment and sites labeling were accomplished by PyMol.

## 3. Fluctuation Test

### 3.1. Optimize the scale of training

- ② Start LB cultures of each E. coli strain and grow culturesovernight at 37°C with orbital shaking at 170 rpm.

- ② Prepare selective media.
- ② Dilute the bacteria until the culturedoes not produce colonies on the resistance plate.
- ② Determine the cell density of theovernight culture as precisely as possible , and thetotal number of cells at this time is N0.
- ② Depending on the measured bacterialdensity, the overnight culture was diluted to 30 ml with a complete medium withdifferent glucose concentrations (0.1%, 0.05%, 0.01%, 0.005%) so that each 30 mlsolution contained a total of 100,000 cells.
- ② Measure the cell density of the fourconditions to determine dilution.
- ② Divided into two 96-well plates (48holes per condition), each with 30ul.
- ② Incubate two plates (37°C) at the right temperature and do not shake.
- ② Obtain cell density for ten of the 48holes under each condition to ensure that cell counts vary with glucoseconcentration.
- ② The remaining holes add 70ul distilledwater to the 30ul culture and mix well before applying the plate to prevent thebottom residue. The coated plate remains dry.
- ② Calculate the number of colonies ofzero-level events and find an appropriate proportion of zero-level events (unmutated plates/total plates between 10% and 80%)

### 3.2. Fluctuation test

- ② Overnight culture the bacteria liquid.
- ② Break up the bacteria liquid and measurecell density.
- ② Dilute the bacteria into a fullycultured solution at the selected glucose concentration.

### 3.3. Apply a separate strain

- ② Confirm the dilution multiple, thenremove a certain volume of the bacteria fluid from the diluted bacteria to coatthe common medium, and then count the colonies (which can be used as theinitial number of cells in each hole).

- ② Add the diluted bacteria to 4 96-wellplates.
- ② Two of the 96-well plates were cultured, and the other two were coated in the same selected medium, where only mutated bacteria could grow, and then count the colonies (the number of bacteria that have mutated- “a”)
- ② After the culture of the two 96-wellplates, ten of the holes were extracted to measure the cultured bacterial density, and the total number of cultured bacteria ( $N_f$ ) was calculated. Then the ten holes of bacteria liquid coated into a complete medium, calculate the number of colonies, at this time the number of colonies is the actual number of active bacteria (b), so as to obtain the survival rate of bacteria viability.
- ② Spread all remaining bacteria into the same selected medium and count (number of mutant cells- “c”).

#### 3.4. Data analysis:

Calculate mutation rates:

$$u =$$

$$p_0 =$$

$$viability =$$

[1] Halperin, S.O., Tou, C.J., Wong, E.B. et al. CRISPR-guided DNA polymerases enable diversification of all nucleotides in a tunable window. *Nature* 560, 248–252 (2018). doi: 10.1038/s41586-018-0384-8

[2] Wang H H , Isaacs F J, Carr P A , et al. Programming cells by multiplex genome engineering and accelerated evolution. *Nature* , 460(7257) : 894 -898,2009.

[3] Luan G, Cai z, Li Y, et al. Genome replication engineering assisted continuous evolution (GREACE) to improve microbial tolerance for biofuels production. *Biotechnology for Biofuels*,6(1):137,2013.

[4] Tou, Connor J.; Schaffer, David V.; Dueber, John E. (2020). Targeted diversification in the *S. cerevisiae* genome with CRISPR-guided DNA polymerase I. *ACS Synthetic Biology*, 0, acssynbio.0c00149–.

doi:10.1021/acssynbio.0c00149

[5] Long, Mengfei; Xu, Meijuan et.al (2020). Directed evolution of ornithine cyclodeaminase using an EvolvR-based growth-coupling strategy for efficient biosynthesis of L-proline. *ACS Synthetic Biology*, acssynbio.0c00198–.

doi:10.1021/acssynbio.0c00198

[6] Tinikul, R., Chunthaboon, P., Phonbuppha, J., & Paladkong, T. (2020). Bacterial luciferase: Molecular mechanisms and applications. *The Enzymes*. doi:10.1016/bs.enz.2020.06.001

[7] Reis, Alexander C.; Halper, Sean M.; Vezeau, Grace E(2019). Simultaneous repression of multiple bacterial genes using nonrepetitive extra-long sgRNA arrays. *Nature Biotechnology*, doi:10.1038/s41587-019-0286-9

[8] Campbell, Zachary T.; Weichsel, Andrzej; Montfort, William R.; Baldwin, Thomas O. (2009). Crystal Structure of the Bacterial Luciferase/Flavin Complex Provides Insight into the Function of the  $\beta$  Subunit. *Biochemistry*, 48(26), 6085–6094.

[9] Thompson, T. B. (1996). The 1.5-A Resolution Crystal Structure of Bacterial Luciferase in Low Salt Conditions. *Journal of Biological Chemistry*, 271(36), 21956–21968.

[10] Abu-Soud, H. M., Clark, A. C., Francisco, W. A., Baldwin, T. O., and Raushel, F. M. (1993) Kinetic destabilization of the hydroperoxy flavin intermediate by site-directed modification of the reactive thiol in bacterial luciferase. *J. Biol. Chem.* 268, 7699–7706.]

[11] Haeussler, M., Schönig, K., Eckert, H. et al.

Evaluation of off-target and on-target scoring algorithms and integration into the guide RNA selection tool CRISPOR. *Genome Biol* 17, 148 (2016). <https://doi.org/10.1186/s13059-016-1012-2>

[12] Concordet, Jean-Paul; Haeussler, Maximilian (2018). CRISPOR: intuitive guide selection for CRISPR/Cas9 genome editing experiments and screens. *Nucleic Acids Research*, 0, -. doi:10.1093/nar/gky354

[13] Fluctuation analysis  
<https://fangman-brewer.genetics.washington.edu/fluctuation.html#optimization>

[14] Stapley, J., Feulner, P. G., Johnston, S. E., Santure, A. W. & Smadja, C. M. Variation in recombination frequency and distribution across eukaryotes: patterns and processes. *Phil. Trans. R. Soc. B* 372, 20160455 (2017).

[15] Vos, M. & Didelot, X. A comparison of homologous recombination rates in bacteria and archaea. *ISME J.* 3, 199 (2009)

[16] Phonbuppha J, Tinikul R, Chaiyen P. Use of Bacterial Luciferase as a Reporter Gene in Eukaryotic Systems. *Methods Mol Biol.* 2021;2274:53-65. doi: 10.1007/978-1-0716-1258-3\_6. PMID: 34050462.