PLOS ONE

Evaluation of a library of *loxP* variants with a wide range of recombination efficiencies by Cre

Yuji Yamauchi[1,2], Hidenori Matsukura[1], Keisuke Motone [3], Mitsuyoshi Ueda[1], Wataru Aoki [1]*

1 Division of Applied Life Sciences, Graduate School of Agriculture, Kyoto University, Sakyo-ku, Kyoto, Japan, **2** Japan Society for the Promotion of Science, Sakyo-ku, Kyoto, Japan, **3** Paul G. Allen School of Computer Science and Engineering, University of Washington, Seattle, Washington, United States of America

* aoki.wataru.6a@kyoto-u.ac.jp

Abstract

Sparse labeling of individual cells is an important approach in neuroscience and many other fields of research. Various methods have been developed to sparsely label only a small population of cells; however, there is no simple and reproducible strategy for managing the probability of sparse labeling at desired levels. Here, we aimed to develop a novel methodology based on the Cre-*lox* system to regulate sparseness at desired levels, and we purely analyzed cleavage efficiencies of *loxP* mutants by Cre. We hypothesized that mutations in the *loxP* sequence reduce the recognition efficiency by Cre, which enables the regulation of the sparseness level of gene expression. In this research, we mutagenized the *loxP* sequence and analyzed a library of *loxP* variants. We evaluated more than 1000 mutant *loxP* sequences, including mutants with reduced excision efficiencies by Cre ranging from 0.51% to 59%. This result suggests that these mutant *loxP* sequences can be useful in regulating the sparseness of genetic labeling at desired levels.

Introduction

Sparse labeling is a genetic methodology that is used to label only a small number of cells in an overall population. Sparse labeling has impacted a variety of fields, and is especially important in neuroscience, because a massive diversity of neurons with unique morphologies is present in the nervous system [1, 2] and a tremendous number of neuronal cells exist in the brain; approximately 86 billion in the human brain and 100 million in the mouse brain [3, 4]. In addition, brains are tightly packed with neurons, and their mixed dendritic and axonal processes hamper the visualization of distinct morphologies. The paradigm in which one examines the characteristics of stochastically selected subsets of cells of the same type is particularly useful because it enables single-cell analysis to elucidate the functional logic of neuronal circuits. In this context, there is a great demand for stochastic gene expression for small populations of cells.

Several methodologies have been developed to achieve sparse labeling. One method consists in the use of an animal line with the desired expression patterns after screening transgenic lines with variegated gene expression [5–8]. Other methods have relied on site-specific recombinases (SSRs). In the methods using SSRs, the sparseness level can be regulated in several ways. First, the sparseness level can be regulated by a suitable tamoxifen dosage in a CreER-*lox*-mediated recombination system [9–15]. Second, viral injections of a low titer into Cre driver lines can be implemented [16, 17]. However, these methodologies have major problems in that it is difficult to determine the sparseness levels a priori with reproducibility, because they require very sophisticated experimental techniques or the time-consuming titration of chemical or gene induction conditions to limit the spatial and/or temporal expression of a recombinase.

Methodologies that can regulate sparseness at a predicted level with high reproducibility would be highly useful. In this situation, several methods have been developed in recent years. First, the mononucleotide repeat frameshift (MORF) method is a Cre-dependent sparse cell labeling approach based on mononucleotide repeat frameshifting as a stochastic translational switch. MORF can regulate sparseness levels with high reproducibility [18, 19]. The labeling rate is approximately ~1%– 5%, depending on the targeted cell type and Cre line used. Second, the mosaic analysis with a repressible cell marker (MARCM) and mosaic analysis with double markers (MADM) transgenic approaches have been established to sparsely label cells based on Cre-*lox*-mediated interchromosomal recombination, which occurs during mitosis [20–22]. The labeling rate is approximately ~1%– 5%, depending on the targeted cell type and Cre line used. These methods can label effectors with high reproducibility; however, the labeling rate is dependent on cell type and Cre line and cannot tune sparseness to a desired level. In addition, MADM and MARCM can be used only in cells undergoing mitosis. Third, sparse predictive activity through recombinase competition (SPARC) is a method that utilizes PhiC31 recombinase and two competing *attP* target sequences and an *attB* target sequence [23]. SPARC used three types of progressively truncated *attP* sequences to regulate the sparseness level. SPARC-D (Dense) labeled 48%– 51% of cells, SPARC-I (Intermediate) labeled 17%– 22% of cells, and SPARC-S (Sparse) labeled 3%– 7% of cells. Finally, stochastic gene activation with genetically regulated sparseness (STARS) [24, 25] was derived from Brainbow [26]. The Brainbow system is a method to stochastically label cells using two mutually exclusive *lox* sequences. STARS regulates the sparseness level of the effector by lengthening a spacer DNA sequence between *lox2272* sequences, to regulate the excision efficiency by Cre. STARS transgenes that possess various lengths of spacers can regulate the sparseness level from 5% to 50% of the cell population. These methods yielded highly reproducible sparse labeling; however, it remains difficult to regulate the sparseness rate at desired levels. For example, SPARC can only adjust the sparseness level by three levels. STARS require a very long spacer DNA (e.g., > 10 kb) to achieve low stochastic labeling rate, which hampers the construction of transformants.

To overcome the difficulties of low reproducibility and low regulatability reported previously, we sought to develop a novel methodology that allows achieving sparseness at desired levels with high reproducibility by adapting the widely used Cre-*lox* recombination system [27–31]. In particular, we focused on the Brainbow system mentioned above. We hypothesized that the expression rate of a gene could be regulated by introducing a mutation in one of the *lox* sequences (*lox2272* or *loxP*) to reduce the recognition rate by Cre [32–43]. We performed random mutagenesis on the *loxP* sequence, and obtained mutants with reduced excision rates by Cre. The *loxP* mutants acquired in this study could be used to label genes in a simple and reproducible way and potentially regulate gene expression in cell populations with a desired level of sparseness. This method is likely to be particularly effective because Cre can be applied to a wide range of organisms, including mice, flies, and worms, and can be employed in post-mitotic cells [44–48].

Results

Strategy for achieving recombination efficiency at the desired rate

Cre-mediated recombination occurs between one of the two identical pairs of *lox* sites (a pair of *loxP* sequences and a pair of *lox2272* sequences) in a mutually exclusive way in the Brainbow system [26]. Moreover, excision by one recombination event removes a *lox* site, which is required for the other recombination event to occur. In the genetic circuit in which two pairs of *lox* sequences are inserted alternately and gene A is inserted between *loxP* sequences and gene B is inserted between *lox2272* sequences, the decision between the expression of gene A and gene B becomes stochastic and mutually exclusive (**Fig 1A**). The recognition efficiency of two pairs of *lox* sequences (*loxP* sequences and *lox2272* sequences) is comparable [26]. Thus, the expression rates of gene A and gene B are expected to be approximately the same. To develop a method that can stochastically activate gene expression with a desired sparseness level, we first considered the reaction kinetics of Cre-*lox*-mediated intrachromosomal recombination. We hypothesized that the affinity of Cre for the mutagenized *loxP* sequence could be reduced relative to that for the *lox2272* sequence (**Fig 1B**) and that we could regulate the expression rate of gene A and gene B using the mutagenized *loxP* sequences.

Strategies for the construction and evaluation of a library of mutant *loxP* sequences

We designed a strategy to analyze the effect of a mutation in the *loxP* sequences on the excision rate by Cre in a high-throughput manner (**Fig 2**). First, we mutagenized one of the recombinase binding elements (RBEs) of the *loxP* sequence by PCR and constructed a library of mutant *loxP* sequences (**Fig 2A and 2B**). Next, the library was cloned into a centromere-type plasmid and introduced into *Saccharomyces cerevisiae with a CreEBD system, which is a* β-estradiol-inducible Cre expression construct [49]. Cre was then induced in yeast with the mutant *loxP* library (**Fig 2C**). After Cre induction, the library of mutant *loxP* sequences was extracted from the yeast (**Fig 2D**). Subsequently, the library of mutant *loxP* sequences was subjected to Illumina sequencing and the cleavage rate between *loxP* sequences was quantified (**Fig 2E**). Finally, to verify the accuracy of the Illumina sequencing results, we randomly selected *loxP* variants and quantified their cleavage rates by quantitative polymerase chain reaction (qPCR) (**Fig 2F**).

Design of a library of mutant *loxP* sequences

In Cre-*lox* recombination, Cre forms a complex with *lox* sequences by recognizing inverted repeats consisting of 13 bp on each side of the *lox* sequences, named RBEs [31]. In our study, we mutagenized 13 bp (5′– ATAACTTCGTATA–3′) of the right RBE of the *loxP* sequence. We predicted that an increase in the number of substitutions would result in a reduction of the affinity of *loxP* variants for Cre. Regarding the substitutions of 3 or more bases, it was difficult to obtain sufficient coverage rate because of the exponentially increasing number of combinations. Thus, the goal of this research was to evaluate most mutants with one or two nucleotide substitutions and a fraction of the mutants with substitutions of three or more nucleotides (**Fig 3A**). The right RBE of the *loxP* sequence was designed to preserve the original base at 84.7% probability, to obtain as many 2-base substitutions as possible. For example, if the original base was A, it was designed so that 84.7% remained as A, 5.1% as T, 5.1% as G, 5.1% as C. When the mutated rate in randomized primer is 15.3%, 2-base substitutions are most efficiently obtained. In this case, about 30% of all mutants of *loxP* sequences would have 2 base substitutions (**Fig 3B**). The RBE sequences of 30 samples after mutagenesis by PCR were
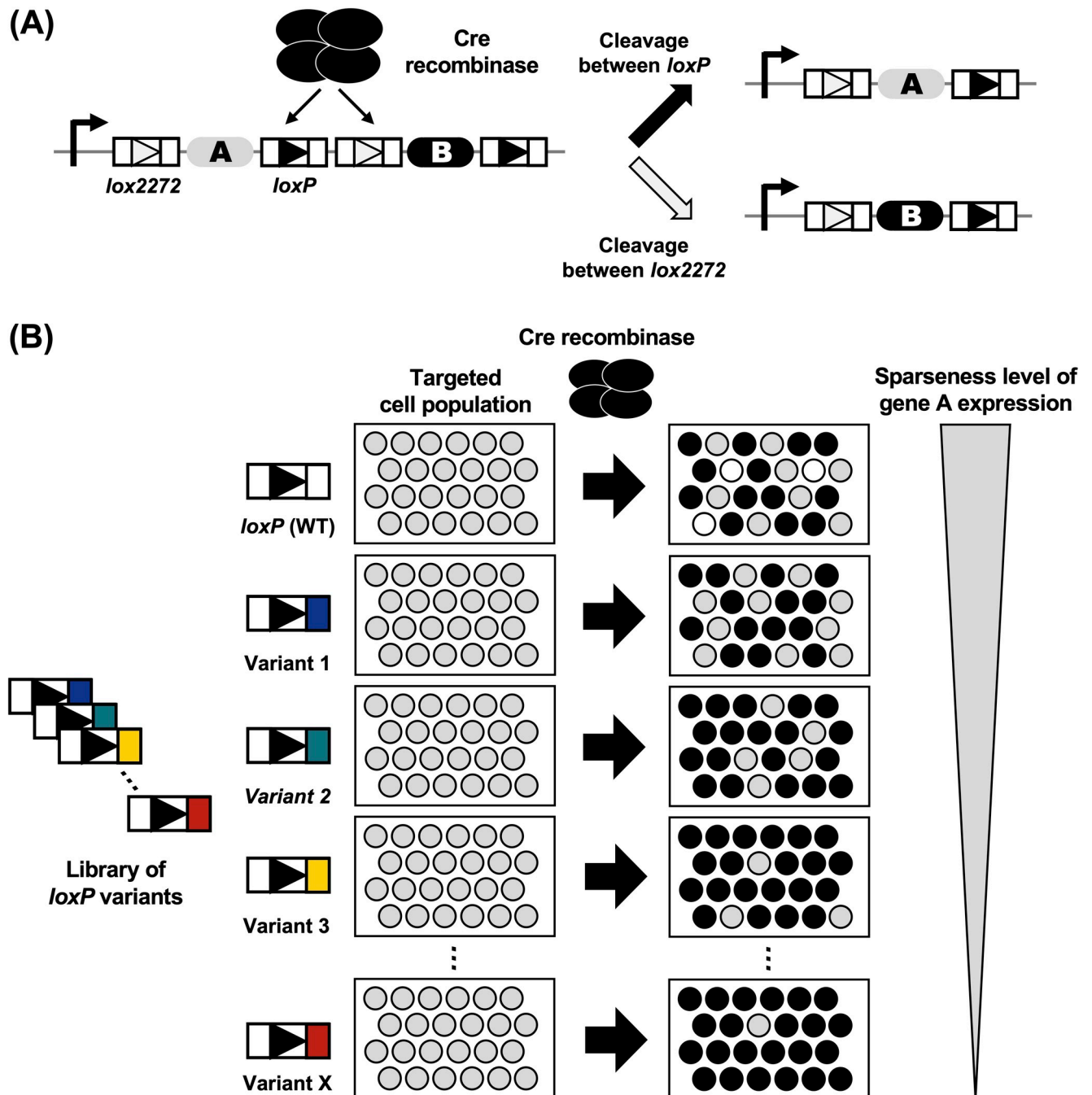
Fig 1. Schematic overview of our strategy. (A) Stochastic excision of *lox* sequences by Cre. Gene A is excised when Cre cleaves *lox2272* sequences and gene B is excised when Cre cleaves *loxP* sequences. Cre cleaves one of two pairs of *lox* sequences (a pair of *lox2272* sequences or a pair of *loxP* sequences) in a mutually exclusive way, which leads to the exclusive expression of gene A or gene B. **(B)** Regulation of the sparse labeling rate using *loxP* variants. We hypothesized that mutations in *loxP* sequences affect the affinity of the *loxP* sequence against Cre, thus rendering them less likely to be cleaved by Cre.

confirmed by Sanger sequencing, and the mutations were successfully introduced at the desired position (**S1 Table**). As a result of the comparison of Sanger sequencing and simulation, we obtained the library of *loxP* variants with approximately the expected rate of the number of substitutions (**Fig 3B**).
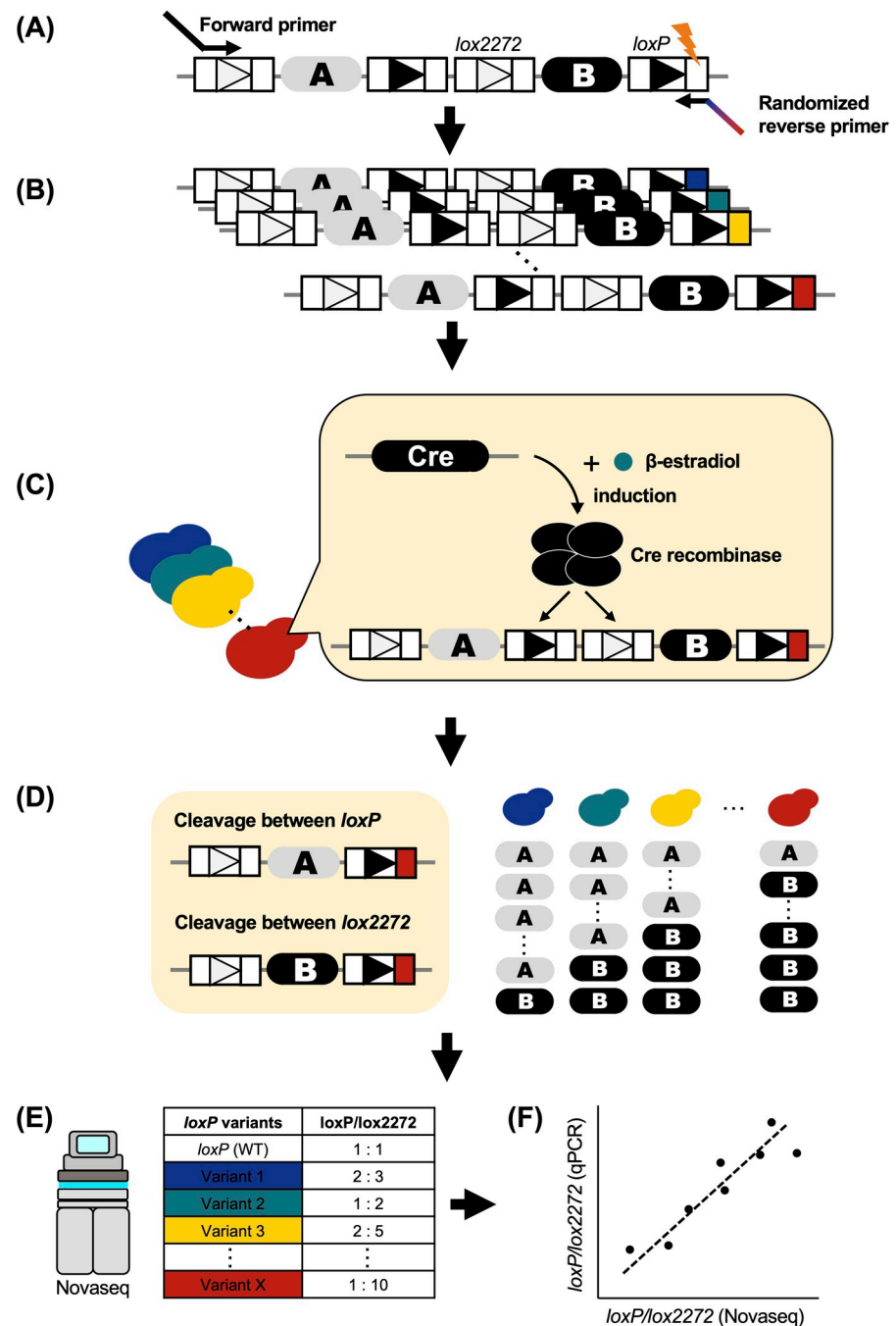
Fig 2. Workflow of this research. Our workflow was divided into six steps. **(A,B) Mutagenesis by PCR.** We mutagenized one of the RBEs of the *loxP* sequence and established a library of mutant *loxP* sequences. **(C) Transformation and Cre induction.** The library of mutant *loxP* sequences was transformed into Saccharomyces cerevisiae and Cre was induced. **(D) Extraction.** After Cre induction, we extracted a library of mutant *loxP* sequences from S. cerevisiae. **(E) Analysis.** We analyzed the excision rate of the mutant *loxP* sequences by Novaseq. **(F) Validation.** We quantified the excision rate of *loxP* variants by qPCR, to validate the results of the Illumina sequencing analysis.

https://doi.org/10.1371/journal.pone.0276657.g002

Analysis of the library of mutant *loxP* sequences by Illumina sequencing

The library of mutagenized *loxP* sequences was analyzed by Illumina sequencing after a cleavage reaction that was carried out by inducing Cre using the CreEBD system in yeast. The
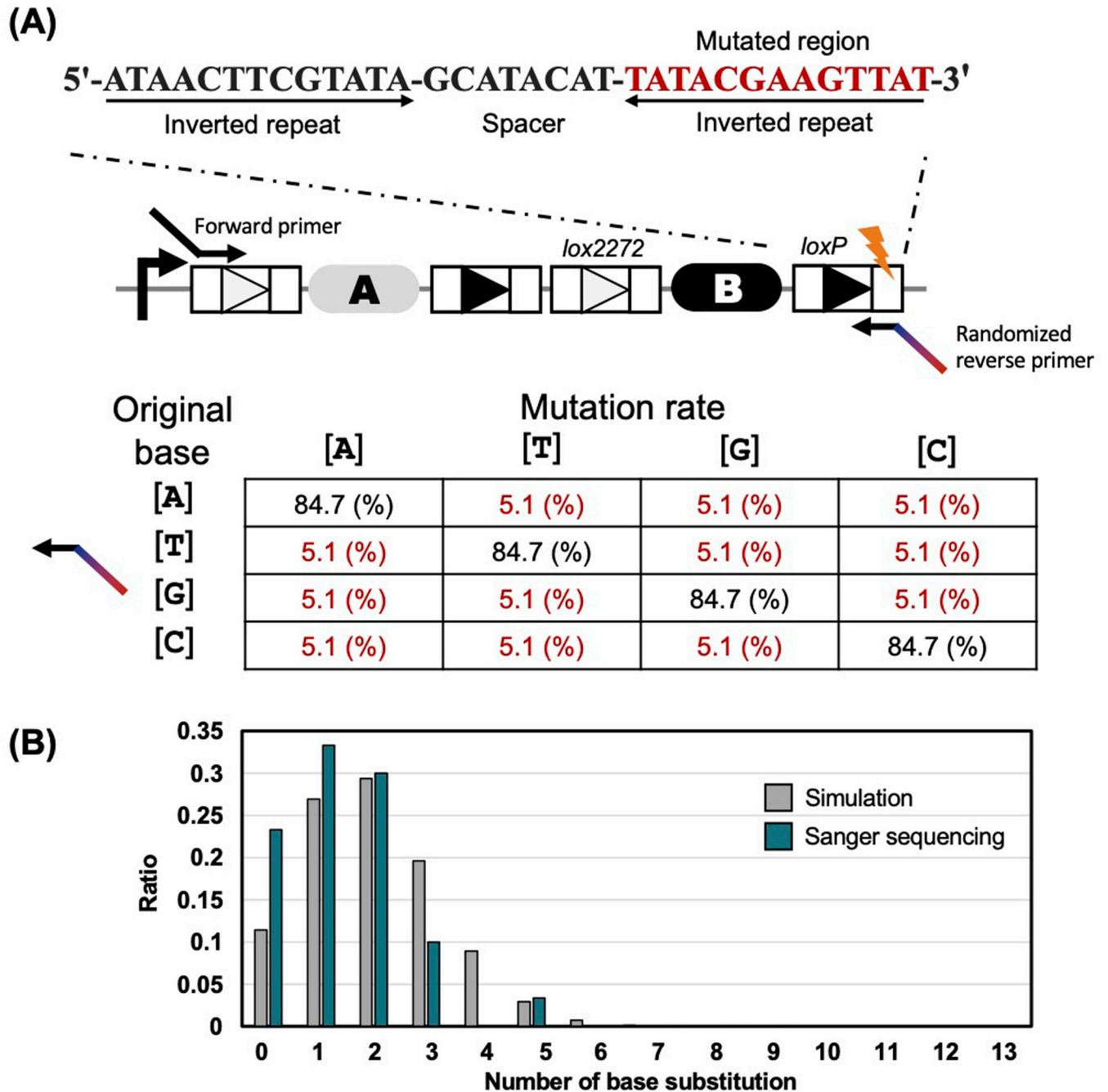
Fig 3. Primer design to construct a library of *loxP* variants. (A) Primer design to mutagenize the *loxP* sequence. We introduced a mutation into the right arm (13 bp) of the *loxP* sequence. The right arm of the *loxP* sequence is shown in red. We set the mutation rate of the primer to 15.3% to obtain as many 2-base substitutions as possible. When the mutated rate in randomized primer is 15.3%, 2-base substitutions are most efficiently obtained. (B) Validation of the distribution of the number of base substitutions by Sanger sequencing. Gray indicates the theoretical value (mutation rate = 15.3%). Green indicates the results of Sanger sequencing. The detailed results of Sanger sequencing are provided in **S1 Table**.

workflow of this analysis is shown in **Fig 4A**. First, Illumina sequencing reads were classified according to the type of mutation of the *loxP* sequences (**Fig 4A1 and 4A2**). In total, we generated 16471 different mutants of *loxP* sequences. To obtain accurate data, we focused on 1111 *loxP* variants with a number of Illumina sequencing reads greater than 500 (**S1 File**). Next, we
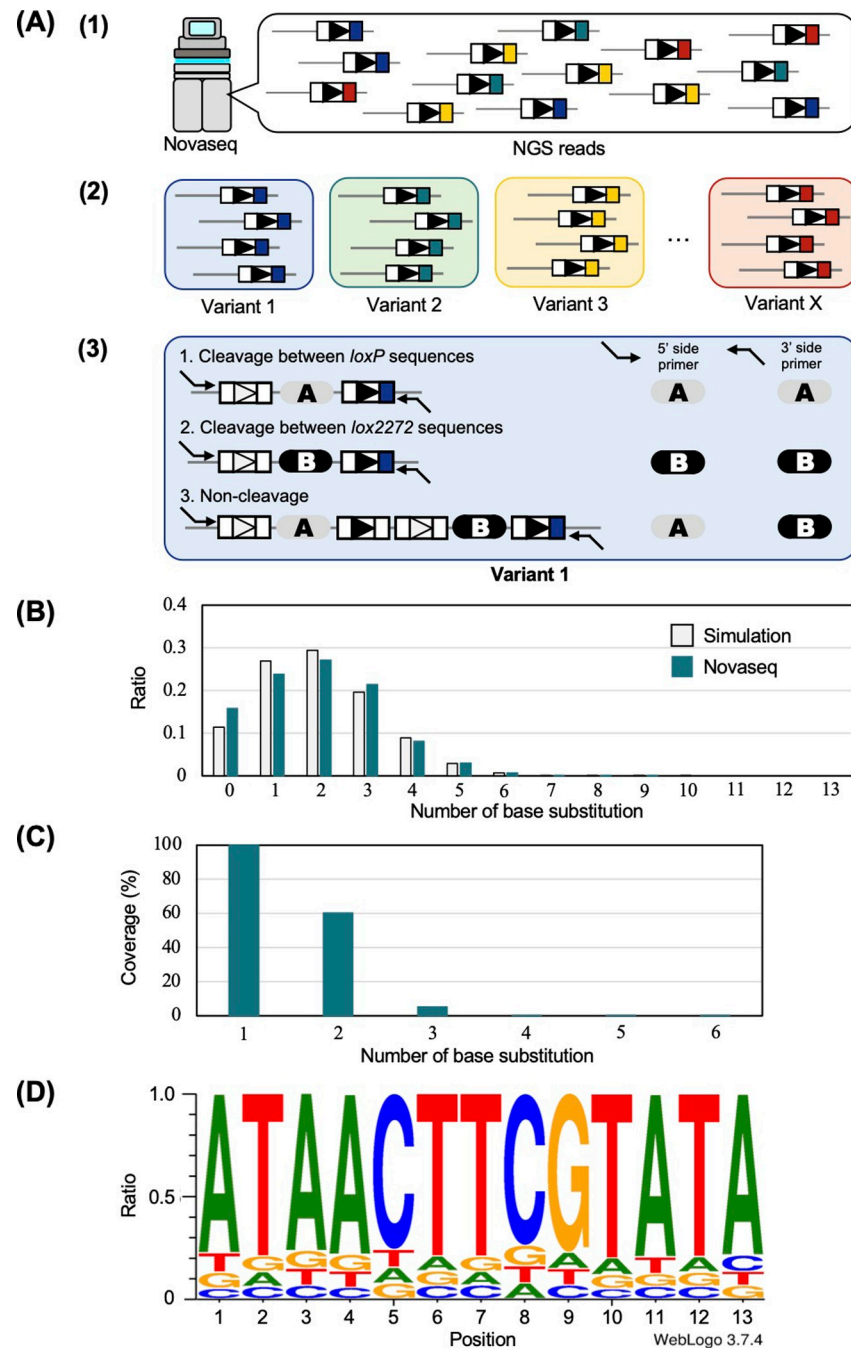
Fig 4. Analysis of the cleavage rate of *loxP* variants by Novaseq. (A) Schematic overview of the Illumina sequencing analysis. We analyzed the Illumina sequencing data in three steps. (1) We extracted DNA sequences that possessed sequences of *loxP* variants. (2) All paired-end Illumina sequencing reads were classified according to the type of mutation in *loxP* sequences. (3) Using the paired-end Illumina sequencing reads obtained using the primer at the 5′ side and the primer at the 3′ side, we identified the patterns of recombination of Illumina sequencing reads for each *loxP* variant. If an Illumina sequencing read possessed sequence A, we determined that the plasmid was cleaved between *loxP* sequences by Cre. If an Illumina sequencing read possessed sequence B, we determined that the plasmid was cleaved between *lox2272* sequences by Cre. If an Illumina sequencing read possessed sequences A and B, we determined that the plasmid was not cleaved. **(B)** Comparison of the distribution of the number of substituted bases between experimental values and theoretical values. The light-gray color indicates the theoretical values (substitution rate = 15.3%). Green indicates the experimental value. This method of calculation is provided in the Materials and Methods. **(C)** Coverage rate (%) of substitution. **(D)** Nucleotide rate at each position of *loxP* sequences.

analyzed the cleavage rate of individual *loxP* mutant sequences using the combination of the Illumina sequencing reads acquired using a 5′ primer with those acquired using a 3′ primer (**Fig 4A3**). As a result of the calculation, the average non-cleavage rate was 3.5%, which confirmed that Cre was correctly induced. Before performing a detailed analysis of individual mutant *loxP* sequences, we confirmed the quality of the library of mutagenized *loxP* sequences. First, the distribution pattern of the number of base substitutions roughly corresponded to the theoretical one (**Fig 4B**), as previously suggested in **Fig 3B** using Sanger sequencing. The coverage rate of base substitution was examined. As a result, we acquired 100% (39/39 variants) of the single-base substitutions and 60.5% (425/702 variants) of the 2-base substitutions (**Fig 4C**). Third, we assessed the nucleotide bias at every 13 positions of the right RBE of the *loxP* sequence. The results showed that no bias existed in the type of bases that were introduced into each of the positions of the RBE (**Fig 4D**).

Subsequently, we quantified the cleavage rate of individual *loxP* sequences. As predicted, as the number of substitutions increased from a 1-base substitution to 6-base substitutions, the average cleavage rate between *loxP* sequences decreased monotonically (**Fig 5A**). As shown in **Fig 5B**, we obtained mutant *loxP* sequences with reduced recognition efficiencies by Cre in various proportions (**Table 1**). These results support our hypothesis that sparseness can be regulated by mutagenizing an RBE of *lox* sequences.

Quantification of the cleavage rate of the mutant *loxP* sequences by qPCR

We assessed the cleavage rate of the mutagenized *loxP* sequences by qPCR to validate the accuracy of the results of the Illumina sequencing analysis. The results indicated that the addition of various mutations to the RBE can alter the cleavage efficiency of Cre at various rates. However, the sequencing results potentially have a certain bias because it requires PCR during sample preparation. To confirm that there is no significant bias in the sequencing results, we performed qPCR, which is low-throughput but can accurately quantify the cleavage rates by Cre. We quantified the non-cleavage rate and *loxP* cleavage rate in each mutant *loxP* sequence by qPCR. A comparison of non-cleavage rates in the absence or presence of Cre induction showed that the recombination events that occurred were dependent on Cre induction (**Fig 6A**). The cleavage rates of the randomly selected *loxP* variants were quantified using qPCR and compared with the results of the Illumina sequencing analysis (**Fig 6B** and **S2 Table**). These results showed that the cleavage rate between *loxP* sequences was highly correlated with that obtained by Illumina sequencing analysis ($R^2 = 0.9695$) (**Fig 6C**). This qPCR result showed that the Illumina sequencing data are reliable.

Identification of mutant *loxP* sequences with less than 1% recognition efficiency

Although we analyzed 1111 variants and successfully identified variants with a recognition efficiency ranging from 17% to 59% (**Fig 5B**), the sparse labeling method requires mutant *loxP* sequences with a lower cleavage efficiency as low as 1%. Fig 5A indicates that the recognition rate by Cre drops as the number of base substitutions increase. To obtain mutant *loxP* sequences with lower recognition rates by Cre, we evaluated the recognition rate of mutant *loxP* sequences with more than 7-base substitutions. As a result of qPCR quantification, we obtain several mutant *loxP* sequences with recognition efficiencies of less than 1% (**Fig 7** and **S2 Table**).

Discussion

The problems of sparseness labeling identified in previous studies included the low regulatability of the sparseness level of the effector with low reproducibility. In this study, we developed a novel sparseness labeling method that overcame the weakness of the previous studies. We

(A)

(B)
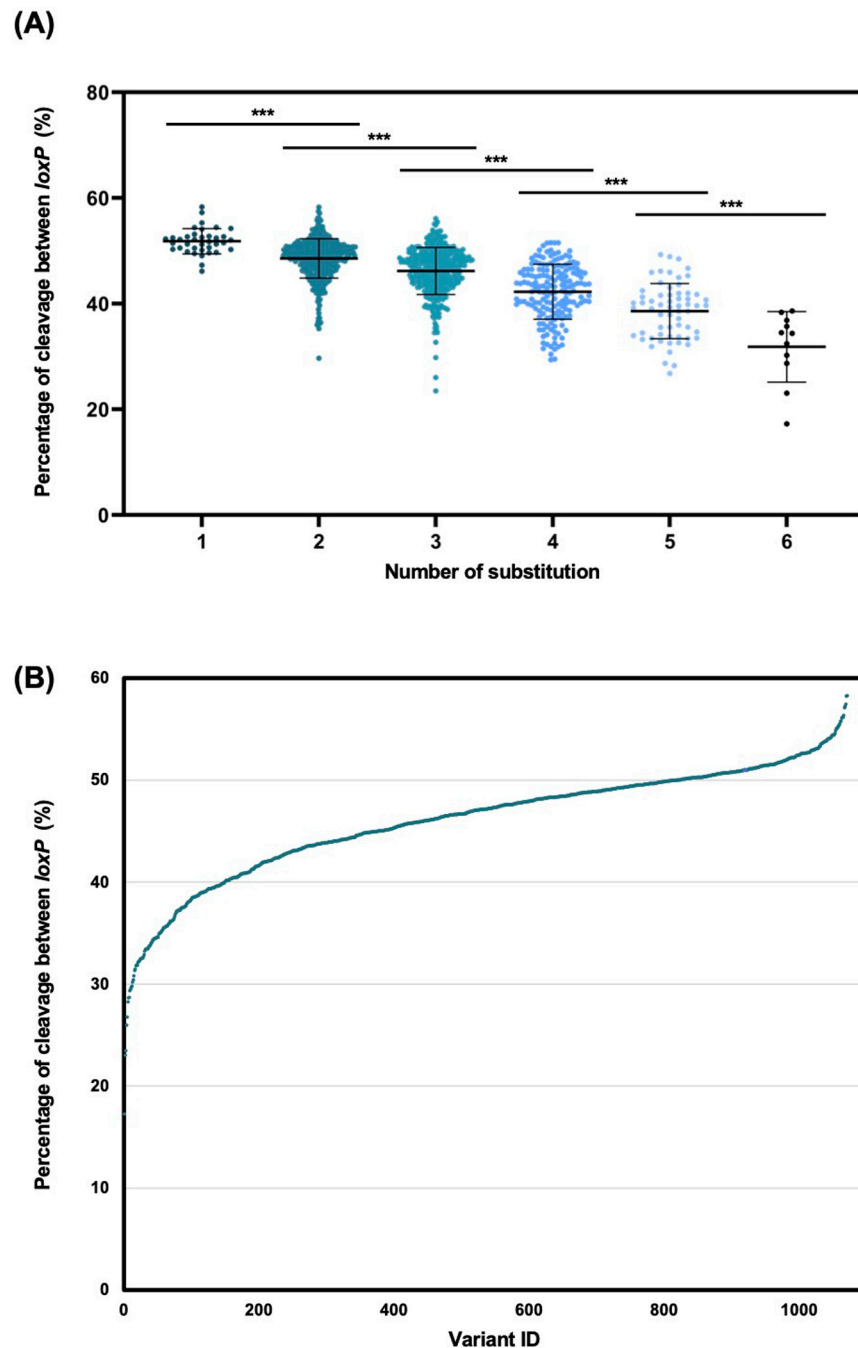
Fig 5. Illumina sequencing data analysis of the *loxP* variants. (**A**) Cleavage rate of all *loxP* variants is classified according to the number of base substitutions. A significant difference was found between each number of substitutions. *** $P < 0.001$, two-tailed *t*-test. (**B**) Distribution of all *loxP* variants. The vertical axis indicates the cleavage rate between *loxP* sequences. *loxP* variants with reduced recognition rate by Cre were listed in order of the cleavage rate between *loxP* sequences. There was some overlap between the plots.

https://doi.org/10.1371/journal.pone.0276657.g005

aimed to obtain mutant *loxP* sequences by introducing random mutations into the right RBE sequence according to our hypothesis that *loxP* variants with reduced recognition efficiency by Cre can regulate the sparseness level desirably. We analyzed 1111 variants and successfully identified variants with a recognition efficiency ranging from 17% to 59% (**Fig 5B**).

Table 1. List of top 50 mutant *loxP* sequences with the lowest recognition rate by Cre among evaluated *loxP* variants. The underlined bases indicate mutated positions.

Sequence of *loxP* variants			Number of substitutions	Cleavage rate between *loxP* sequences (%)
Left REB	**Spacer**	**Right RBE**		
ATAACTTCGTATA	-GCATACAT-	GATGTCAAGATAG	6	17.3
ATAACTTCGTATA	-GCATACAT-	TGGAGCATGTCAT	6	23
ATAACTTCGTATA	-GCATACAT-	CGTACAAAGTTAT	3	23.5
ATAACTTCGTATA	-GCATACAT-	TCAACCAAGTTAT	3	26
ATAACTTCGTATA	-GCATACAT-	TGCAACAAGTCAT	5	26.8
ATAACTTCGTATA	-GCATACAT-	TCAACACATTTAT	5	28.3
ATAACTTCGTATA	-GCATACAT-	GATACTTATTGAC	6	28.7
ATAACTTCGTATA	-GCATACAT-	CATACCTCGTCAT	5	28.7
ATAACTTCGTATA	-GCATACAT-	AAATCGAAGTCAT	4	29.4
ATAACTTCGTATA	-GCATACAT-	AAAGCGGAGTTAT	4	29.5
ATAACTTCGTATA	-GCATACAT-	AAAACGAAGTTAT	2	29.7
ATAACTTCGTATA	-GCATACAT-	AAAACCAAGTTAT	3	29.8
ATAACTTCGTATA	-GCATACAT-	TTCACCCAGTTTC	6	30.2
ATAACTTCGTATA	-GCATACAT-	TCAACCAAGTAAT	4	30.4
ATAACTTCGTATA	-GCATACAT-	CTGACGGAGTGAT	5	30.8
ATAACTTCGTATA	-GCATACAT-	AAAACGATGTAAT	4	31.3
ATAACTTCGTATA	-GCATACAT-	TCTACATAGATAT	4	31.5
ATAACTTCGTATA	-GCATACAT-	TACATGACGCTAT	4	31.8
ATAACTTCGTATA	-GCATACAT-	TACATCACCTTAT	5	31.9
ATAACTTCGTATA	-GCATACAT-	TACATAAAGTCAT	4	31.9
ATAACTTCGTATA	-GCATACAT-	TACACTGAGTTAG	4	32.2
ATAACTTCGTATA	-GCATACAT-	CATCCCAAACTAT	5	32.2
ATAACTTCGTATA	-GCATACAT-	GATAAGACGTTGT	4	32.3
ATAACTTCGTATA	-GCATACAT-	GATTAGATGCTAC	6	32.4
ATAACTTCGTATA	-GCATACAT-	TGTAGGCAGATAG	5	32.5
ATAACTTCGTATA	-GCATACAT-	GATATGCAGTCAT	4	32.5
ATAACTTCGTATA	-GCATACAT-	TACACAAATTCAT	4	32.5
ATAACTTCGTATA	-GCATACAT-	GATAAAAAATTAC	5	32.6
ATAACTTCGTATA	-GCATACAT-	CATATGACGTTAT	3	32.7
ATAACTTCGTATA	-GCATACAT-	GATATCACGTTCT	5	32.9
ATAACTTCGTATA	-GCATACAT-	GGTACGCACTTCT	5	33.2
ATAACTTCGTATA	-GCATACAT-	CTTAGCCAGTTAT	5	33.4
ATAACTTCGTATA	-GCATACAT-	TTCACGGGATTAT	5	33.4
ATAACTTCGTATA	-GCATACAT-	GACACCAAGATAT	4	33.5
ATAACTTCGTATA	-GCATACAT-	TACACGGAACTAG	5	33.5
ATAACTTCGTATA	-GCATACAT-	TGTACTAAGTTCG	4	33.5
ATAACTTCGTATA	-GCATACAT-	TCTACTATTTTAA	5	33.7
ATAACTTCGTATA	-GCATACAT-	TGACCGACGTTAG	5	33.7
ATAACTTCGTATA	-GCATACAT-	CATACCAAATTCT	4	33.8
ATAACTTCGTATA	-GCATACAT-	TACACGTTATTCT	5	34
ATAACTTCGTATA	-GCATACAT-	CTTACCAAGTTTT	4	34.1
ATAACTTCGTATA	-GCATACAT-	TTTGTCACGTTAT	5	34.2
ATAACTTCGTATA	-GCATACAT-	TTCTGGAAGTTAT	4	34.3
ATAACTTCGTATA	-GCATACAT-	CGAGCGAACCTAT	6	34.4
ATAACTTCGTATA	-GCATACAT-	CATTCAAAAGTAT	5	34.5

(Continued)

Table 1. (Continued)

Sequence of *loxP* variants			Number of substitutions	Cleavage rate between *loxP* sequences (%)
Left REB	**Spacer**	**Right RBE**		
ATAACTTCGTATA	-GCATACAT-	GATATGAAATTAT	3	34.5
ATAACTTCGTATA	-GCATACAT-	CATATGTTGTCAA	6	34.5
ATAACTTCGTATA	-GCATACAT-	TGTACGGAGTTCT	3	34.5
ATAACTTCGTATA	-GCATACAT-	GATACCAAGTCAT	3	34.6
ATAACTTCGTATA	-GCATACAT-	CATCGGAATTTAA	5	34.6

This result supports our hypothesis that the efficiency of recognition of the *loxP* sequence by Cre can be regulated by precisely introducing mutations into the arm of the *lox* sequence. However, the mutated *loxP* sequences maintained a higher recognition efficiency by Cre than initially expected. Even in the case of three or more base substitutions, the recognition efficiency of *loxP* variants remained relatively high. If we adopt sparse labeling in dense tissues, such as the brain, a mutant that can achieve a high sparsity with a labeling rate of 1% or less is needed. As shown in **Fig 5A**, our experiments confirmed that the efficiency of recognition of *loxP* by Cre decreased as the number of substitutions introduced into the RBE of the *loxP* sequence increases. As shown in Fig 7, we evaluated the recognition rate of mutant *loxP* sequences with more than 7-base substitutions, and we obtained several mutant *loxP* sequences with recognition efficiencies of less than 1%. We also found that some mutants with fewer base substitutions may have lower cleavage rates than mutants with more base substitutions, as shown in Fig 7. This result indicates that the evaluation of individual mutant *loxP* sequences is vital to achieving sparse labeling at any desired rates.

As mentioned in the introduction, several previous studies evaluated the effect of mutations on Cre recombinase or *loxP* sequences [32–43]. Hartung, M. & Kisters-Woike, B. evaluated the effect of the mutation in Cre recombinase [41]. However, it is difficult to precisely regulate the sparseness level by introducing mutations into Cre recombinase. On the other hand, our approach can control the sparseness level only by selecting a mutant *loxP* sequence. Also, we can use existing Cre lines. Missirlis, P. I., *et al.* introduced mutations in the spacer region of the *loxP* sequences [34]. The spacer sequence determines the specificity. Hence, we think that many of the mutants would result in a complete loss of recognition by Cre. Thus, introducing mutations into the spacer region is not an appropriate approach to regulate the sparseness level. Sheren J *et al.* examined the effect of mutations on the spacer region and RBE of the *loxP* sequence, respectively. The method of evaluating mutant *loxP* sequences in our experiment is very different from that of Sheren J *et al.* In our study, the *lox2272* sequence competes with the mutant *loxP* sequence. In contrast, Sheren J *et al.* evaluated mutant *loxP* sequences alone. In our experiment, Cre cannot clave the other *lox* sequence if Cre cleaves one *lox* sequence. On the other hand, if the mutant *loxP* sequence is present alone, the Cre can always cleave the *loxP* sequence while Cre is acting. Therefore, the cleavage rate of mutant *loxP* sequences measured by Sheren J *et al.* under competitive conditions with other *lox* sequences such as *lox2272* is unclear. In addition, we have evaluated cleavage rates for over 1000 mutant *loxP* sequences in this study, allowing us to adjust sparse labeling rates very strictly ranging from 0.51%–59%. In summary, this research is the first large dataset for measuring the cleavage rate of mutant *loxP* sequences in competitive conditions with other *lox* sequences (*lox2272* sequence in this study). Thus, our dataset may allow for regulating sparseness levels at the desired rate.

This study provides the proof of concept for establishing a novel method that allows sparse labeling at desired rates. The novel sparse labeling method proposed in this study has the
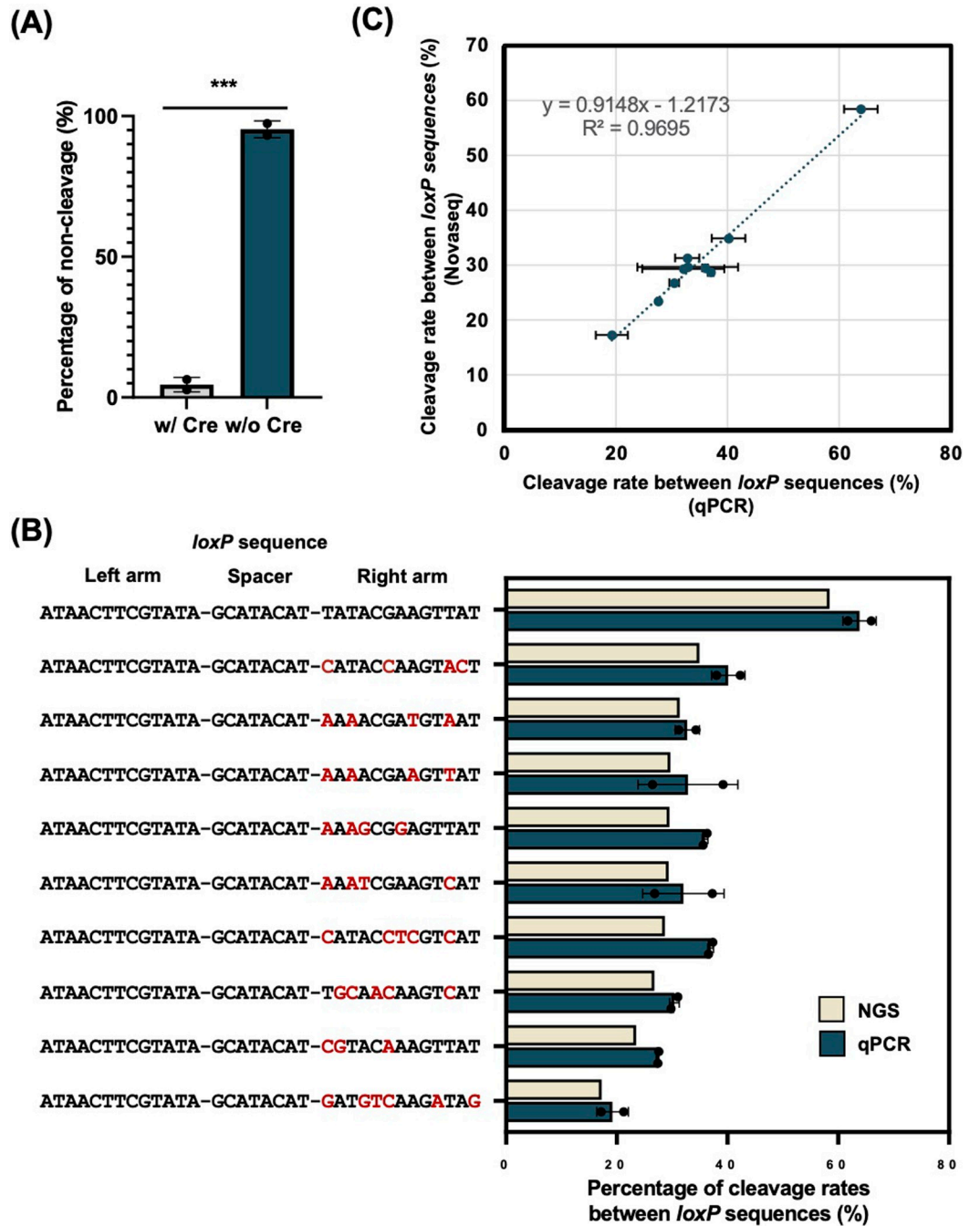
Fig 6. Quantification of the cleavage rate of *loxP* variants by qPCR. (A) Comparison of non-cleavage rates in the presence or absence of Cre induction. The error bars represent the standard deviation. A significant difference was found in the presence or absence of Cre induction (two-tailed *t*-test, N = 3). *** $P < 0.001$ **(B)** Comparison of the results of the Illumina sequencing analysis and qPCR. We randomly selected nine *loxP* variants and the WT *loxP* sequence for qPCR. We quantified the cleavage rate of the *loxP* sequence in each variant and WT (see details in the Materials and Methods). The mutated nucleotide is shown in red. The bar graph in light yellow shows the cleavage rate of *loxP* sequences quantified by Illumina sequencing, whereas green indicates the results of qPCR. Each plot was averaged across two independent experiments. The error bars represent the standard deviation. **(C)** Correlation of the cleavage rate between *loxP* sequences, as quantified by Illumina sequencing analysis and qPCR. Each plot was averaged across two independent experiments. The error bars represent the standard deviation.
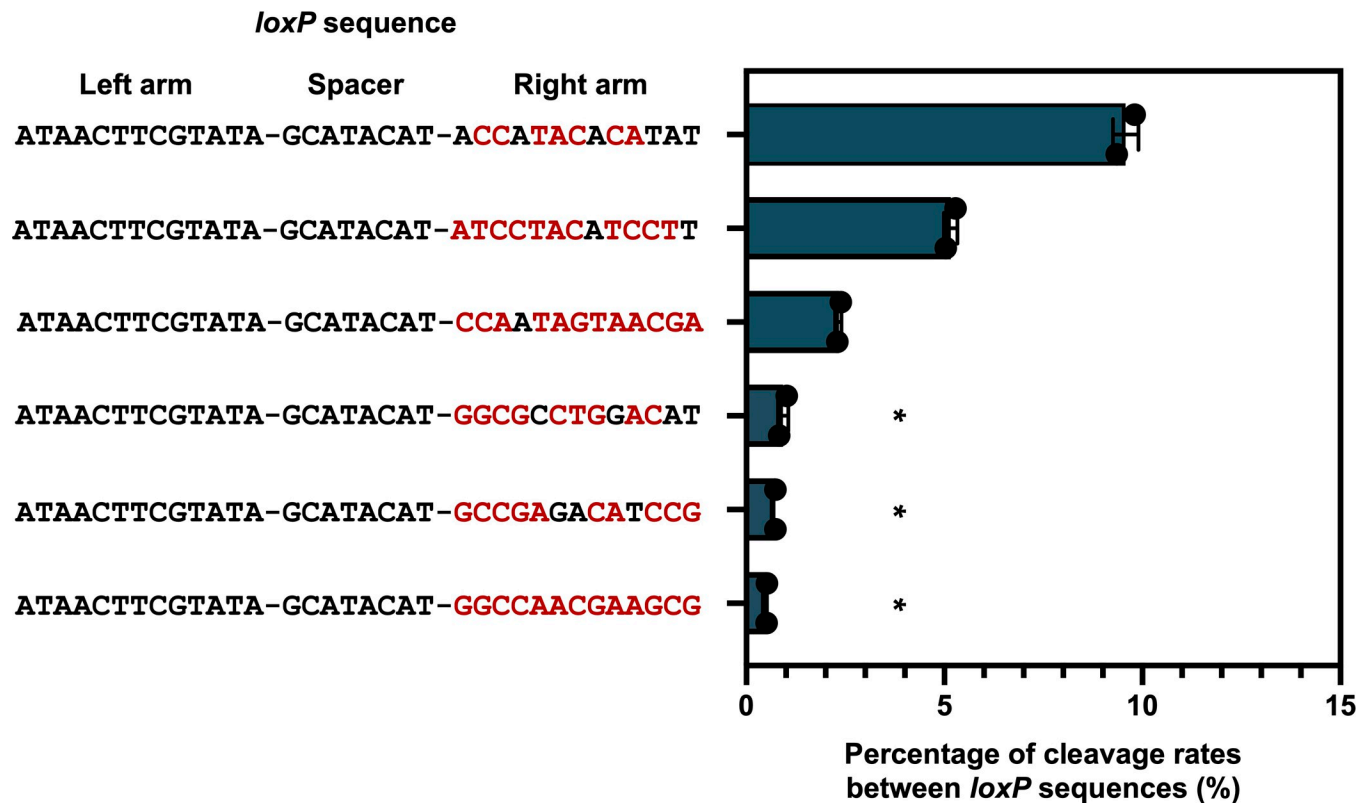
loxP sequence

Left arm	Spacer	Right arm

ATAACTTCGTATA–GCATACAT–ACCATACACATAT

ATAACTTCGTATA–GCATACAT–ATCCTACATCCTT

ATAACTTCGTATA–GCATACAT–CCAATAGTAACGA

ATAACTTCGTATA–GCATACAT–GGCGCCTGGACAT

ATAACTTCGTATA–GCATACAT–GCCGAGACATCCG

ATAACTTCGTATA–GCATACAT–GGCCAACGAAGCG

**Percentage of cleavage rates
between loxP sequences (%)**

Fig 7. The quantification of loxP variants with more than 7-base substitutions by qPCR. We randomly selected *loxP* mutants with more than 7-base substitutions. We quantified the cleavage rate of the each mutant *loxP* sequence using qPCR. The mutated nucleotide is shown in red. Mutants marked with an asterisk showed less than 1% of recognition rate by Cre. The bar graph in green indicates the results of qPCR. Each plot was averaged across two independent experiments. The error bars represent the standard deviation.

https://doi.org/10.1371/journal.pone.0276657.g007

potential to overcome some of the disadvantages of previous methods. The screening of transgenic lines with desired expression patterns [5–8], titration of a suitable tamoxifen dosage [9–15], or amount of viral injection [16, 17] require very sophisticated experimental techniques or the time-consuming titration of chemical or gene induction conditions to limit the spatial and temporal expression of a recombinase. In our method, screening or titrating the chemical and genetic induction conditions is unnecessary to achieve the desired sparseness because a mutant *loxP* sequence that can achieve a desired labeling rate is selected in advance. The MORF and MADM/MARCM methods have a disadvantage that the sparseness level cannot be controlled [18–22]. The SPARC and STARS methods have the problem of low controllable sparseness levels [23, 24]. On the other hand, our method can regulate the sparseness level in more than 1000 patterns ranging from 0.51% to 59% by appropriate mutant *loxP* sequence from the mutant *loxP* library.

The ability to regulate the sparseness level at the desired rate with high reproducibility is a major advantage of this method. One example is the combination of our method with the method that uses a cellomics approach [48]. The cellomics approach method applies the Cre-*lox* system to stochastically label opsin in a small population of neural networks. If the mutant *loxP* sequences can be used in the cellomics approach, we could intervene in the activity of a smaller number of neurons. In other examples, our method can be applied to generate a genetic mosaic for analyzing the population dosage of genes involved in sporadic genetic illnesses, as well as to promote cancer development. By applying this method to stepwise change

normal cells into cells expressing cancer-inducing genes in a cell population at desired levels, it may be possible to mimic the environment of cell competition when a minority population of cancer cells is present in the majority of the normal cell population [50–55]. Taken together, as a purely genetic method, it has the potential to be adapted to a variety of fields of research.

Finally, we discuss the limitation of our sparse labeling method. The primary purpose of this study is to present data on the cleavage rate of mutant *loxP* sequencing, which is a proof-of-concept for establishing a novel methodology that allows regulating sparseness levels at the desired rate. We have not yet tested our methodology in neural tissues, and we also have not yet tested the labeling rate in different cell types. As a future experimental plan, it is necessary to test our method in several types of neural tissue and validate the possibility of regulating sparseness levels at the desired rates. Another potential limitation also exists. When combining the mutant *loxP* sequences with the Brainbow system, we can use existing mouse lines for the Cre expression system. For the Brainbow transgenic lines, we have to establish a new line for each mutant *loxP* sequence. There is currently no way to reduce this effort, and it will need to be solved in the future.

Materials and methods

Construction of a library of mutant *loxP* sequences

A library of mutant *loxP* sequences was constructed via PCR using the pRS416-*lox* plasmid as a template (**S2 File**). We mutagenized the *loxP* sequences using primer_No. 1 and primer_No. 2, as shown in **S3 Table**. The primers were synthesized commercially via solid-phase synthesis (Eurofin, Tokyo, Japan). The underlined 13-base sequence in primer_No. 2 indicates the mutagenized location of the *loxP* sequence. The underlined bases were synthesized using biased randomization with an 84.7% chance of retaining the original sequence (e.g., A means A: 84.7%, T: 5.1%, G: 5.1%, C: 5.1%). This is because this study aimed to evaluate *loxP* mutants with one or two nucleotide substitutions and a certain extent of mutants with three or more nucleotide substitutions. When $x = 84.7$, the two nucleotide substitutions are most efficiently obtained. We calculated the distribution of the number of base substitutions using the following formula:

$$\frac{13!}{(y!(13-y)!)} \times \left(\frac{x}{100}\right)^{13-y} \times \left(\frac{100-x}{100}\right)^{y}$$

In this formula, the retention rate [%] is represented by x and the number of base substitutions is represented by y.

The library was transformed in competent *E. coli* DH5α (F⁻, Φ80d*lacZΔ*M15, *Δ*(*lacZYA-argF*)U169, *deoR*, *recA*1, *endA*1, *hsdR*17(r$_K^-$, m$_K^+$), *phoA*, *supE*44, λ⁻, *thi*-1, *gyrA*96, *relA*1). The transformed *E. coli* was cultured in a Luria-Bertani (LB) medium (1% [w/v] tryptone, 0.5% [w/v] yeast extract and 1% [w/v] sodium chloride) containing 100 μg/mL ampicillin. Then, plasmids were extracted using the FastGene Plasmid Minikit (NIPPON Genetics, Tokyo, Japan, FG-90502).

Cre induction in yeast

We constructed β-estradiol-inducible Cre-expressing yeast. The *S. cerevisiae* strain BY4741 (*MATa*, *his3Δ1*, *leu2Δ0*, *met15Δ0*, *ura3Δ0*) was used as the host. The pRS403-CreEBD plasmid (**S2 File**) was inserted genomically into the yeast *his3Δ1* site. Using a Frozen EZ Yeast Transformation II Kit (Zymo Research, Irvine, CA, USA, T2001), yeast cells were transformed with the constructed plasmid. The transformants were screened on a synthetic defined (SD) solid

medium without L-histidine. The components of the solid SD medium were 0.67% [w/v] yeast nitrogen base without amino acids, 2% [w/v] glucose, and 2% [w/v] agar with appropriate amino acids and a nucleobase (0.012% [w/v] L-leucine, 0.002% [w/v] L-methionine, and 0.002% [w/v] uracil). The mutagenized *loxP* library was transformed into a yeast strain possessing the CreEBD transgene. The transformants were screened on an SD solid medium without L-histidine and uracil. A total of more than 10,000 colonies were collected to prevent the loss of diversity in the *loxP* library. For 24 h, the colonies were precultured on liquid SD medium at 30°C and 250 rpm. The precultured yeast was inoculated into 10 mL of liquid SD medium containing D-galactose as a carbon source with an OD_{600} of 1. Yeast cells were cultured at 30°C and 250 rpm for 24 h.

Preparation of DNA samples for Illumina sequencing analysis

The *loxP* library after Cre induction was extracted using the Zymoprep^{TM} Yeast Plasmid Miniprep II kit (ZYMO RESEARCH, D2004). The extracted samples were amplified by PCR using primer_No. 3 and primer_No. 4. The cycling parameters were as follows: 94°C for 2 min; followed by five cycles at 98°C for 10 s/55°C for 5 s/68°C for 30 s, 68°C for 7 min; and final hold at 4°C. The number of PCR cycles was set to five to reduce PCR bias as much as possible. The prepared samples were sequenced using Novaseq6000 (Illumina, San Diego, CA, USA) at the paired end. The quality check of samples, the addition of adaptor sequences, the addition of index sequences, and Illumina sequencing runs were performed using the services of MacroGen Japan (Tokyo, Japan).

Illumina sequencing data analysis

The Python script shown in **S3 File** was used to analyze the Illumina sequencing data. Briefly, using fastq files and the Python script, the sequences of *loxP* variants were extracted and the number of reads was counted. We set the cut-off for analysis at 500 to obtain reliable *loxP* cleavage rates. We classified the cleavage patterns of each Illumina sequencing read (non-cleavage, cleavage between *loxP*, and cleavage between *lox2272*) using the paired-end sequence information. We calculated the cleavage rate between *loxP* sequences based on the following formula:

$$\frac{\text{The number of NGS reads cleaved between } loxP}{\text{The number of NGS reads cleaved between } loxP \text{ or } lox2272} \times 100$$

The processed data are presented in **S1 File**.

Quantitative polymerase chain reaction (qPCR)

To confirm the accuracy of the results of the Illumina sequencing analysis, we conducted a validation experiment by quantitative polymerase chain reaction (qPCR). Nine *loxP* variants were selected randomly from sequences with a decreased cleavage rate compared with WT *loxP*. We constructed plasmids that possessed the target mutation via PCR using pRS416_Leu_Ura (**S2 File**) as a template. The primers used to amplify the DNA fragments are shown in **S3 Table** (primer_No. 5 to No. 15). Then, we prepared three sets of primers for qPCR: (1) primers to quantify the cleavage rate between *lox2272* (primer_No. 16 and primer_No. 17), (2) primers to quantify the cleavage rate between *loxP* (primer_No. 18 and primer_No. 19), and (3) primers to quantify the non-cleavage rate (primer_No. 20 and primer_No. 21). In addition, we used pRS416_Leu_Ura, pRS416_Leu_Ura_ΔloxP, and pRS416_Leu_Ura_Δlox2272 (**S2 File**) for generating a calibration curve. A dilution series of 7 points in 10-fold increments from 1.0×10^7 copies/μL to 1 copy/μL was used for the calibration

curve. The non-cleavage rate, the cleavage rate between *loxP*, and the cleavage rate between *lox2272* of individual mutant *loxP* sequences were calculated using these calibration curves. The PCR mixture included: 17 μL of distilled water, 25 μL of FastStart SYBR Green Master (without ROX) (Roche, Basel, Switzerland, 04673484001), 1.5 μL of 10 pmol/μL forward primer, 1.5 μL of 10 pmol/μL reverse primer, and 5 μL of the template plasmid. PCR was carried out on a StepOnePlus™ instrument (Thermo Fisher Scientific, USA) using the following cycling conditions for all primer sets: 95˚C for 10 min; followed by 40 cycles of 95˚C for 15s, 60˚C for 30s; and 1 cycle of 95˚C for 15s, 60˚C for 1 min, and 95˚C for 10 s.

Supporting information

S1 Table. Sequences of *loxP* variants confirmed by Sanger sequencing.
(XLSX)

S2 Table. List of *loxP* variants measured by qPCR.
(XLSX)

S3 Table. Primers used in this study.
(XLSX)

S1 File. List of all *loxP* variants evaluated in this research.
(XLSX)

S2 File. Plasmid maps used in this study. The full sequences of the plasmids used in this study are shown.
(XLSX)

S3 File. Python scripts for data analysis.
(TXT)

Author Contributions

Conceptualization: Yuji Yamauchi, Wataru Aoki.

Data curation: Yuji Yamauchi, Hidenori Matsukura, Wataru Aoki.

Formal analysis: Yuji Yamauchi.

Funding acquisition: Yuji Yamauchi, Wataru Aoki.

Investigation: Yuji Yamauchi, Hidenori Matsukura.

Methodology: Yuji Yamauchi, Hidenori Matsukura, Wataru Aoki.

Project administration: Yuji Yamauchi, Wataru Aoki.

Resources: Yuji Yamauchi, Hidenori Matsukura.

Software: Yuji Yamauchi, Keisuke Motone.

Supervision: Mitsuyoshi Ueda, Wataru Aoki.

Validation: Yuji Yamauchi, Hidenori Matsukura.

Visualization: Yuji Yamauchi.

Writing – original draft: Yuji Yamauchi, Wataru Aoki.

Writing – review & editing: Yuji Yamauchi, Hidenori Matsukura, Keisuke Motone, Mitsuyoshi Ueda, Wataru Aoki.

References

1. Tasic B, Menon V, Nguyen TN, Kim TK, Jarsky T, Yao Z, et al. Adult mouse cortical cell taxonomy revealed by single cell transcriptomics. Nat Neurosci 2016 192. 2016; 19: 335–346. https://doi.org/10.1038/nn.4216 PMID: 26727548

2. Swanson LW. Quest for the basic plan of nervous system circuitry. Brain Res Rev. 2007; 55: 356–372. https://doi.org/10.1016/j.brainresrev.2006.12.006 PMID: 17267046

3. Herculano-Houzel S. The remarkable, yet not extraordinary, human brain as a scaled-up primate brain and its associated cost. Proc Natl Acad Sci U S A. 2012; 109: 10661–10668. https://doi.org/10.1073/pnas.1201895109 PMID: 22723358

4. Herculano-Houzel S, Mota B, Lent R. Cellular scaling rules for rodent brains. Proc Natl Acad Sci U S A. 2006; 103: 12138–12143. https://doi.org/10.1073/pnas.0604911103 PMID: 16880386

5. Xiao T, Roeser T, Staub W, Baier H. A GFP-based genetic screen reveals mutations that disrupt the architecture of the zebrafish retinotectal projection. Development. 2005; 132: 2955–2967. https://doi.org/10.1242/dev.01861 PMID: 15930106

6. Feng G, Mellor RH, Bernstein M, Keller-Peck C, Nguyen QT, Wallace M, et al. Imaging neuronal subsets in transgenic mice expressing multiple spectral variants of GFP. Neuron. 2000; 28: 41–51. https://doi.org/10.1016/s0896-6273(00)00084-2 PMID: 11086982

7. Young P, Qiu L, Wang D, Zhao S, Gross J, Feng G. Single-neuron labeling with inducible Cre-mediated knockout in transgenic mice. Nat Neurosci 2008 116. 2008; 11: 721–728. https://doi.org/10.1038/nn.2118 PMID: 18454144

8. Choong SC, Rosenberg SS, Fancy SPJ, Zhao C, Shen YAA, Hahn AT, et al. Neurite outgrowth inhibitor Nogo-A establishes spatial segregation and extent of oligodendrocyte myelination. Proc Natl Acad Sci U S A. 2012; 109: 1299–1304. https://doi.org/10.1073/pnas.1113540109 PMID: 22160722

9. Hayashi S, McMahon AP. Efficient recombination in diverse tissues by a tamoxifen-inducible form of Cre: A tool for temporally regulated gene activation/inactivation in the mouse. Dev Biol. 2002; 244: 305–318. https://doi.org/10.1006/dbio.2002.0597 PMID: 11944939

10. Badea TC, Wang Y, Nathans J. A Noninvasive genetic/pharmacologic strategy for visualizing cell morphology and clonal relationships in the mouse. J Neurosci. 2003; 23: 2314–2322. https://doi.org/10.1523/JNEUROSCI.23-06-02314.2003 PMID: 12657690

11. Rotolo T, Smallwood PM, Williams J, Nathans J. Genetically-directed, cell type-specific sparse labeling for the analysis of neuronal morphology. PLoS One. 2008; 3: e4099. https://doi.org/10.1371/journal.pone.0004099 PMID: 19116659

12. Ellisor D, Zervas M. Tamoxifen dose response and conditional cell marking: Is there regurate? Mol Cell Neurosci. 2010; 45: 132–138. https://doi.org/10.1016/J.MCN.2010.06.004 PMID: 20600933

13. Gómez-Casati ME, Murtie J, Taylor B, Corfas G. Cell-specific inducible gene recombination in postnatal inner ear supporting cells and glia. JARO J Assoc Res Otolaryngol. 2010; 11: 19. https://doi.org/10.1007/s10162-009-0191-x PMID: 19820996

14. Wu H, Williams J, Nathans J. Morphologic diversity of cutaneous sensory afferents revealed by genetically directed sparse labeling. Elife. 2012; 2012. https://doi.org/10.7554/eLife.00181 PMID: 23256042

15. Badea TC, Hua ZL, Smallwood PM, Williams J, Rotolo T, Ye X, et al. New mouse lines for the analysis of neuronal morphology using CreER(T)/loxP-directed sparse labeling. PLoS One. 2009; 4: e7859. https://doi.org/10.1371/journal.pone.0007859 PMID: 19924248

16. Lin R, Wang R, Yuan J, Feng Q, Zhou Y, Zeng S, et al. Cell-type-specific and projection-specific brainwide reconstruction of single neurons. Nat Methods 2018 1512. 2018; 15: 1033–1036. https://doi.org/10.1038/s41592-018-0184-y PMID: 30455464

17. Zingg B, Peng B, Huang J, Tao HW, Zhang LI. Synaptic specificity and application of anterograde transsynaptic AAV for probing neural circuitry. J Neurosci. 2020; 40: 3250–3267. https://doi.org/10.1523/JNEUROSCI.2158-19.2020 PMID: 32198185

18. Lu XH, Yang XW. Genetically-directed sparse neuronal labeling in BAC transgenic mice through mononucleotide repeat frameshift. Sci Reports 2017 71. 2017; 7: 1–10. https://doi.org/10.1038/srep43915 PMID: 28272512

19. Veldman MB, Park CS, Eyermann CM, Zhang JY, Zuniga-Sanchez E, Hirano AA, et al. Brainwide genetic sparse cell labeling to illuminate the morphology of neurons and glia with Cre-dependent MORF mice. Neuron. 2020; 108: 111-127.e6. https://doi.org/10.1016/j.neuron.2020.07.019 PMID: 32795398

20. Luo L. Fly MARCM and mouse MADM: Genetic methods of labeling and manipulating single neurons. Brain Res Rev. 2007; 55: 220–227. https://doi.org/10.1016/J.BRAINRESREV.2007.01.012 PMID: 17408568

21. Zong H, Espinosa JS, Su HH, Muzumdar MD, Luo L. Mosaic analysis with double markers in mice. Cell. 2005; 121: 479–492. https://doi.org/10.1016/j.cell.2005.02.012 PMID: 15882628

22. Lee T, Luo L. Mosaic analysis with a repressible cell marker for studies of gene function in neuronal morphogenesis. Neuron. 1999; 22: 451–461. https://doi.org/10.1016/s0896-6273(00)80701-1 PMID: 10197526

23. Isaacman-Beck J, Paik KC, Wienecke CFR, Yang HH, Fisher YE, Wang IE, et al. SPARC enables genetic manipulation of precise proportions of cells. Nat Neurosci 2020 239. 2020; 23: 1168–1175. https://doi.org/10.1038/s41593-020-0668-9 PMID: 32690967

24. Wang SZ, Liu BH, Tao HW, Xia K, Zhang LI. A Genetic strategy for stochastic gene activation with regulated sparseness (STARS). PLoS One. 2009; 4: 4200. https://doi.org/10.1371/JOURNAL.PONE.0004200 PMID: 19145242

25. Ibrahim LA, Huang JJ, Wang SZ, Kim YJ, Zhang LI, Tao HW. Sparse labeling and neural tracing in brain circuits by STARS strategy: revealing morphological development of type II spiral ganglion neurons. Cereb Cortex. 2021; 31: 2759–2772. https://doi.org/10.1093/CERCOR/BHY154 PMID: 29982390

26. Livet J, Weissman TA, Kang H, Draft RW, Lu J, Bennis RA, et al. Transgenic strategies for combinatorial expression of fluorescent proteins in the nervous system. Nat 2007 4507166. 2007; 450: 56–62. https://doi.org/10.1038/nature06293 PMID: 17972876

27. Sternberg N, Hamilton D. Bacteriophage P1 site-specific recombination: I. Recombination between *loxP* sites. J Mol Biol. 1981; 150: 467–486. https://doi.org/10.1016/0022-2836(81)90375-2

28. Hoess RH, Ziese M, Sternberg N. P1 site-specific recombination: nucleotide sequence of the recombining sites. Proc Natl Acad Sci U S A. 1982; 79: 3398. https://doi.org/10.1073/PNAS.79.11.3398 PMID: 6954485

29. Van Duyne GD. A structural view of Cre-loxP site-specific recombination. 2003; 30: 87–104. https://doi.org/10.1146/ANNUREV.BIOPHYS.30.1.87

30. Van Duyne GD. Cre Recombinase. Microbiol Spectr. 2015;3. https://doi.org/10.1128/microbiolspec.MDNA3-0014-2014 PMID: 26104563

31. Meinke G, Bohm A, Hauber J, Pisabarro MT, Buchholz F. Cre recombinase and other tyrosine recombinases. Chem Rev. 2016; 116: 12785–12820. https://doi.org/10.1021/acs.chemrev.6b00077 PMID: 27163859

32. Lee G, Saito I. Role of nucleotide sequences of *loxP* spacer region in Cre-mediated recombination. Gene. 1998; 216: 55–65. https://doi.org/10.1016/S0378-1119(98)00325-4

33. Thomson JG, Rucker EB, Piedrahita JA. Mutational analysis of *loxP* sites for efficient Cre-mediated insertion into genomic DNA. genesis. 2003; 36: 162–167. https://doi.org/10.1002/GENE.10211 PMID: 12872248

34. Missirlis PI, Smailus DE, Holt RA. A high-throughput screen identifying sequence and promiscuity characteristics of the *loxP* spacer region in Cre-mediated recombination. BMC Genomics. 2006;7. https://doi.org/10.1186/1471-2164-7-73 PMID: 16595017

35. Sheren J, Langer SJ, Leinwand LA. A randomized library approach to identifying functional *lox* site domains for the Cre recombinase. Nucleic Acids Res. 2007; 35: 5464–5473. https://doi.org/10.1093/NAR/GKM604 PMID: 17702764

36. Siegel RW, Jain R, Bradbury A. Using an in vivo phagemid system to identify non-compatible *loxP* sequences. FEBS Lett. 2001; 505: 467–473. https://doi.org/10.1016/S0014-5793(01)02806-X

37. Hoess RH, Wierzbicki A, Abremski K. The role of the *loxP* spacer region in P1 site-specific recombination. Nucleic Acids Res. 1986; 14: 2287. https://doi.org/10.1093/NAR/14.5.2287 PMID: 3457367

38. Sauer B. Multiplex Cre/*lox* recombination permits selective site-specific DNA targeting to both a natural and an engineered site in the yeast genome. Nucleic Acids Res. 1996; 24: 4608–4613. https://doi.org/10.1093/NAR/24.23.4608 PMID: 8972844

39. Albert H, Dale EC, Lee E, Ow DW. Site-specific integraten of DNA into wild-type and mutant *lox* sites placed in the plant genome. Plant J. 1995; 7: 649–659. https://doi.org/10.1046/J.1365-313X.1995.7040649.X PMID: 7742860

40. Langer SJ, Ghafoori AP, Byrd M, Leinwand L. A genetic screen identifies novel non-compatible *loxP* sites. Nucleic Acids Res. 2002; 30: 3067–3077. https://doi.org/10.1093/NAR/GKF421 PMID: 12136089

41. Hartung M, Kisters-Woike B. Cre mutants with altered DNA binding properties. J Biol Chem. 1998; 273: 22884–22891. https://doi.org/10.1074/jbc.273.36.22884 PMID: 9722507

42. Rüfer AW, Sauer B. Non-contact positions impose site selectivity on Cre recombinase. Nucleic Acids Res. 2002; 30: 2764. https://doi.org/10.1093/nar/gkf399 PMID: 12087159

43. Eroshenko N, Church GM. Mutants of Cre recombinase with improved accuracy. Nat Commun 2013 41. 2013; 4: 1–10. https://doi.org/10.1038/ncomms3509 PMID: 24056590

44. Hadjieconomou D, Rotkopf S, Alexandre C, Bell DM, Dickson BJ, Salecker I. Flybow: genetic multicolor cell labeling for neural circuit analysis in Drosophila melanogaster. Nat Methods 2011 83. 2011; 8: 260–266. https://doi.org/10.1038/nmeth.1567 PMID: 21297619

45. Hampel S, Chung P, McKellar CE, Hall D, Looger LL, Simpson JH. Drosophila Brainbow: a recombinase-based fluorescence labeling technique to subdivide neural expression patterns. Nat Methods 2011 83. 2011; 8: 253–259. https://doi.org/10.1038/nmeth.1566 PMID: 21297621

46. Sethi S, Wang JW. A versatile genetic tool for post-translational regurate of gene expression in Drosophila melanogaster. Elife. 2017; 6. https://doi.org/10.7554/ELIFE.30327 PMID: 29140243

47. Kage-Nakadai E, Imae R, Suehiro Y, Yoshina S, Hori S, Mitani S. A conditional knockout toolkit for Caenorhabditis elegans based on the *Cre/loxP* recombination. PLoS One. 2014; 9: e114680. https://doi.org/10.1371/JOURNAL.PONE.0114680 PMID: 25474529

48. Aoki W, Matsukura H, Yamauchi Y, Yokoyama H, Hasegawa K, Shinya R, et al. Cellomics approach for high-throughput functional annotation of Caenorhabditis elegans neural network. Sci Reports 2018 81. 2018; 8: 1–9. https://doi.org/10.1038/s41598-018-28653-x PMID: 29991757

49. Jia B, Wu Y, Li BZ, Mitchell LA, Liu H, Pan S, et al. Precise regurate of SCRaMbLE in synthetic haploid and diploid yeast. Nat Commun 2018 91. 2018; 9: 1–13. https://doi.org/10.1038/s41467-018-03084-4 PMID: 29789567

50. Muzumdar MD, Luo L, Zong H. Modeling sporadic loss of heterozygosity in mice by using mosaic analysis with double markers (MADM). Proc Natl Acad Sci U S A. 2007; 104: 4495–4500. https://doi.org/10.1073/PNAS.0606491104 PMID: 17360552

51. Muzumdar MD, Dorans KJ, Chung KM, Robbins R, Tammela T, Gocheva V, et al. Clonal dynamics following p53 loss of heterozygosity in Kras-driven cancers. Nat Commun 2016 71. 2016; 7: 1–13. https://doi.org/10.1038/ncomms12685 PMID: 27585860

52. Liu C, Sage JC, Miller MR, Verhaak RGW, Hippenmeyer S, Vogel H, et al. Mosaic analysis with double markers reveals tumor cell of origin in glioma. Cell. 2011; 146: 209–221. https://doi.org/10.1016/j.cell.2011.06.014 PMID: 21737130

53. Gonzalez PP, Kim J, Galvao RP, Cruickshanks N, Abounader R, Zong H. p53 and NF 1 loss plays distinct but complementary roles in glioma initiation and progression. Glia. 2018; 66: 999–1015. https://doi.org/10.1002/glia.23297 PMID: 29392777

54. Tian A, Kang B, Li B, Qiu B, Jiang W, Shao F, et al. Oncogenic state and cell identity combinatorially dictate the susceptibility of cells within glioma development hierarchy to IGF1R targeting. Adv Sci. 2020; 7: 2001724. https://doi.org/10.1002/advs.202001724 PMID: 33173731

55. Yao M, Ventura PB, Jiang Y, Rodriguez FJ, Wang L, Perry JSA, et al. Astrocytic trans-differentiation completes a multicellular paracrine feedback loop required for medulloblastoma tumor growth. Cell. 2020; 180: 502-520.e19. https://doi.org/10.1016/j.cell.2019.12.024 PMID: 31983537