MICROBIOLOGY SOCIETY

OPEN ACCESS

Lox'd in translation: contradictions in the nomenclature surrounding common lox-site mutants and their implications in experiments

Daniel Shaw[1], Luis Serrano[1,2,3] and Maria Lluch-Senar[1,4,*]

Abstract

The Cre-Lox system is a highly versatile and powerful DNA recombinase mechanism, mainly used in genetic engineering to insert or remove desired DNA sequences. It is widely utilized across multiple fields of biology, with applications ranging from plants, to mammals, to microbes. A key feature of this system is its ability to allow recombination between mutant lox sites. Two of the most commonly used mutant sites are named lox66 and lox71, which recombine to create a functionally inactive double mutant lox72 site. However, a large portion of the published literature has incorrectly annotated these mutant lox sites, which in turn can lead to difficulties in replication of methods, design of proper vectors and confusion over the proper nomenclature. Here, we demonstrate common errors in annotations, the impacts they can have on experimental viability, and a standardized naming convention. We also show an example of how this incorrect annotation can induce toxic effects in bacteria that lack optimal DNA repair systems, exemplified by *Mycoplasma pneumoniae*.

DATA SUMMARY

The authors confirm all supporting data, code and protocols have been provided within the article or through supplementary data files.

INTRODUCTION

The Cre-Lox system was first characterized in 1981 by Nat Sternberg and Daniel Harrison [1]. It is a DNA recombinase system derived from the P1 bacteriophage, and consists of two components. The first, the Cre recombinase, is a 38 kDa monomeric tyrosine recombinase. This protein acts upon a pair of 34 bp lox sites [**l**ocus **o**f (**x**)crossing over], which consist of an 8 bp central spacer region flanked by two 13 bp inverted repeat regions [2, 3]. The central spacer unit of the lox site gives it a 'directionality', which dictates the outcome of the reaction between the Cre recombinase and DNA, as shown in Fig. 1. If the two lox sites are in the same (*cis*) orientation, recombination results in the circularization of the genomic DNA between the two lox sites, and thus its removal from the genome. If the lox sites are in opposite (*trans*) orientations,

then the DNA between them is inverted as a consequence of the recombination event.

For the sake of brevity, a full review into the history of the Cre recombinase and its reaction kinetics are outside the scope of this paper, and have been discussed in detail by other authors [4–7]. Instead, here we will focus on one aspect of the Cre-Lox system; its ability to utilize mutant lox sites, and how these sites interact.

The two most widely used lox mutations are the lox66 and lox71 sites, which were first described in 1995 [8]. These lox mutants contain alterations to their initial or terminal five base pairs, with the lox66 having the first five bases changed, and lox71 having the last five bases changed (in regard to the canonical loxP), shown below in Table 1.

The Cre recombinase exists as a monomer, however during recombination it forms a tetramer, with one molecule attached to each flanking region of the two lox sites [5]. As shown in Fig. 2 below, modified from the review of the Cre recombinase by Van Duyne [5], the Cre complex
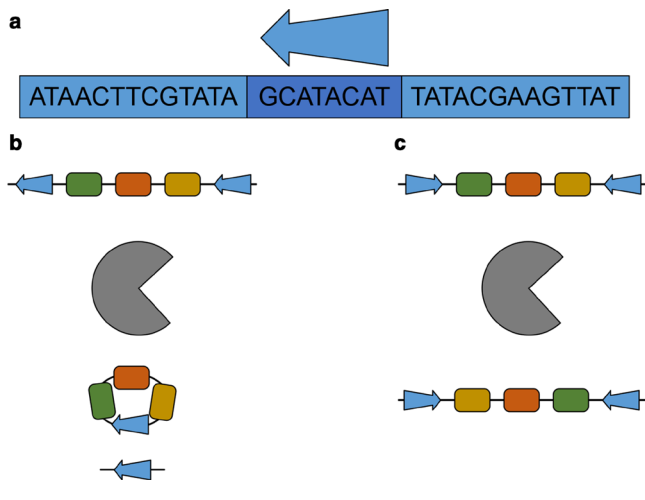
a

b **c**

Fig. 1. Sequence and orientation of the WT loxP site, and effect of orientation on lox recombination reactions. (a) Sequence of the loxP site, with the 13 bp inverted repeats in lighter blue, and the central spacer region in darker blue. The arrow indicates the directionality of the lox site. (b) Action of the Cre recombinase (grey) on two lox sites in *cis* orientation. DNA between the two sites is excised and circularized, leaving one lox site in the genome and one lox site on the circular DNA. (c) Action of the Cre recombinase when lox sites are in *trans*. DNA between the two sites is inverted.

Impact Statement

Cre-Lox recombination is a widely utilised biotechnological application. However, within the current literature there is a small but significant number of papers that display miss-annotations of the common mutant sites lox66 and lox71. Here, we identify what form these errors take, why they may have gone undetected, and propose a standardised naming convention to help prevent these errors in the future. We also show how these miss-annotations have the potential to disrupt experiments, documenting how incorrect annotations can lead to the creation of improper lox mutants. In bacteria with less robust DNA repair mechanisms, such as Mycoplasma pneumoniae, the inclusion of the wrong lox site during Cre recombination can lead to a uniformly lethal phenotype.

manoeuvres the two DNA strands so that the lox sites are brought together. The arrows indicate the directionality of the lox sites, and as the complex is formed they face in opposite directions. The lagging strand from both DNA molecules is then cleaved, and annealed to the other DNA molecule via a Holliday junction-forming reaction. The same then occurs for the other strand, resulting in the left element of the first lox site becoming joined to the right element of the second, and vice versa.

As both lox sites in the diagram are identical, they do not change their sequence via recombination, as every flanking region is identical. However, as each new lox site generated is a combination of both previous sites, when two different mutant lox sites are incorporated, a double mutant is always created, as the action of the Cre recombinase complex to manoeuvre the DNA strands into an antiparallel formation ensures that one 'side' of the reaction will always contain both mutant flanking regions, and thus will combine together [9].

Table 1. Sequence of the canonical loxP site and its mutants, with directionality of the lox site from left to right

Lox site	Sequence (5′ to 3′ leading strand)
LoxP	ATAACTTCGTATA GCATACAT TATACGAAGTTAT
Lox66	ATAACTTCGTATA GCATACAT TATACGAAcggta
Lox71	taccgTTCGTATA GCATACAT TATACGAAGTTAT
Lox72	taccgTTCGTATA GCATACAT TATACGAAcggta

The Cre recombinase can recognize the single mutant lox66 and lox71 with the same affinity as the loxP. However, the double mutant lox72 shows a marked decrease in affinity for the Cre [8], making the site functionally silent. This lack of affinity for the lox72 site allows for multiple lox reactions to occur in a genome in series, without affecting or cross reacting with each other [10–12].

In this paper, we show a comprehensive analysis of the literature associated with lox66 and lox71 sites, with examples of the common errors that have occurred in the nomenclature and a standardized naming convention to help resolve these remaining ambiguities. We also highlight the dangers of such errors, demonstrating that the presence of a single active lox site within a bacterial genome can initiate a lethal phenotype, on par with a double-strand break in the DNA. We use *Mycoplasma pneumoniae* as a model organism, and demonstrate that the lethality associated with the Cre acting on a single lox is analogous to a double-stranded break generated by the meganuclease *I-SceI,* which we also report to be active in *M. pneumoniae* for the first time.

With the demonstration that there occurs mis-annotation of both the mutant arms of the lox sites, and also their directionality, we demonstrate the potentially destructive consequences of when these errors are compounded. This occurred within our own lab, and we hope that this publication may avert the same issues in future experiments, and potentially explain why some previous experiments may have failed.

METHODS

Literature review of lox-site nomenclature

A literature review aiming to capture as many of the papers describing the interactions between lox66 and lox71 sites was performed (see Fig. 3). Many papers were present via multiple search terms, so replicates were removed. The papers were
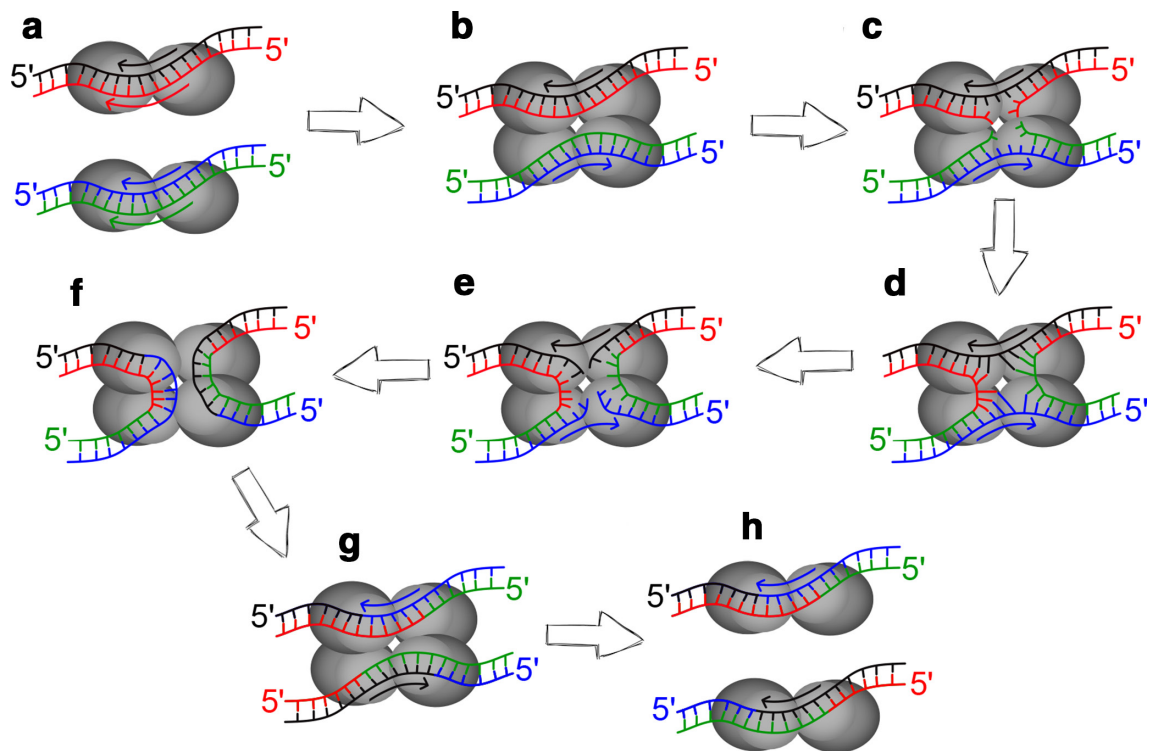
Fig. 2. Modified from Van Duyne [5]. Shows the reaction pathway of the Cre recombinase complex on two loxP sites, with the stands of each DNA molecule shown in separate colours, with the top strands in black and blue, and the bottom strands in red and green. The direction of the lox sites is indicated by the coloured arrows. (a) The Cre recombinase molecules bind to each homology arm of each lox site (b) The Cre-lox complex is formed, with the four Cre molecules bringing the DNA together so the two lox sites run antiparallel to each other. (c) Cleavage of the bottom strands by Tyr324 in the Cre subunits. (d) Formation of 3′-phosphotyrosine linkages between the bottom strands leads to the creation of a four-way Holliday junction between the two DNA molecules. (e) Cleavage of the top strands allows for strand exchange between the two DNA molecules. (f) Top strands of both molecules are ligated together. (g) Formation of the two new lox sites. (h) New lox sites contain the 5′ homology arm of the first lox site, with the 3′ homology arm of the second, and vice versa.

then queried for their relevance to the topic, specifically if they included the use of the lox66 or lox71 site within their methodology or discussion. Those papers that used the sites in a material way were included, and all others discarded.

The remaining 47 papers were then judged for their accuracy in regard to the lox-site nomenclature. The benchmark used for accuracy was the nomenclature described in the original paper by Albert *et al.* [8]. Papers which provided the sequence of the lox sites they used were filtered into either the 'Correct' or 'Incorrect' category, depending on if the nomenclature they used was in concordance with Albert *et al.* All papers that did not show the sequence of the lox sites were placed within the 'Unclear' category.

Strains and culture conditions

Wild-type *M. pneumoniae* strain M129 (ATTC 29342, subtype 1, broth passage no. 35) was used. Cells were cultured in 75 cm³ tissue culture flasks at 37 °C in standard Hayflick media, as described by Hayflick [13] and Yus *et al.*, [14], supplemented with 100 µg ml⁻¹ ampicillin, 2 µg ml⁻¹ tetracycline, 20 µg ml⁻¹ chloramphenicol, 3.3 µg ml⁻¹ puromycin and 200 µg ml⁻¹ gentamicin as appropriate. Hayflick agar plates were

created by supplementing the Hayflick with 1 % Bacto Agar (BD, Cat. No. 214010) pre-autoclave.

NEB 5-alpha Competent *E. coli* cells (New England Biolabs, Cat. No. C2987H) were used for plasmid amplification and cloning. They were grown at 37 °C in Lysogeny Broth (LB) at 200 r.p.m. or static on LB agar plates, supplemented with 100 µg ml⁻¹ ampicillin.

Plasmid DNA

All plasmids were generated using the Gibson isothermal assembly method [15]. DNA was isolated from NEB 5-alpha Competent *E. coli* cells, and individual clones were selected for using LB + ampicillin plates (100 µg ml⁻¹). Correct ligation was confirmed via Sanger sequencing (Eurofins Genomics). A list of all plasmids, and primers used in their generation and sequencing can be found in Tables S1 and S2 (available in the online version of this article).

Single lox-site lethality

WT M129 cells were transformed according to the protocol outlined by Hedreyda *et al.* [16]. Cells were grown to mid-log
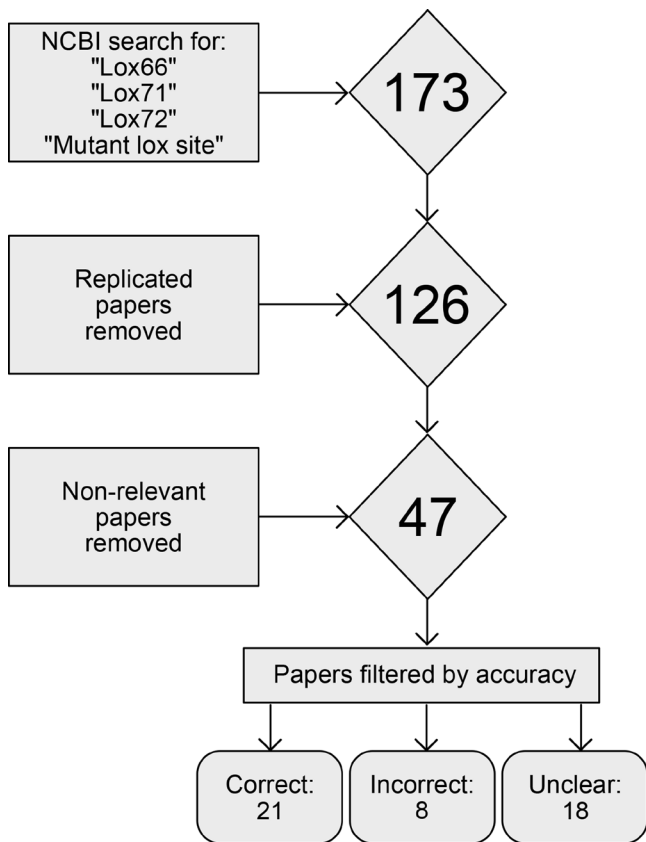
Fig. 3. Flowchart describing the generation of the literature review. The square boxes show the data acquisition and filtering steps in order of occurrence, and the diamonds show the number of papers being considered after each step. The final 47 chosen papers were analysed for their use of lox nomenclature, and classified appropriately, as described in the text above.

phase, identified by the Hayflick media changing from red to orange. The media was decanted and the flask was washed 3× with 10 ml chilled electroporation buffer (EB: 8 mM HEPES, 272 nM sucrose, pH 7.4). Cells were scraped into 500 µl chilled EB and homogenised via 10× passages through a 25-gauge syringe needle. Aliquots of 50 µl of the homogenised cells were mixed with a pre-chilled 30 µl EB solution containing the 1 pMole of pMTn_Lox66_Sce_Cm plasmid DNA. Samples were then kept on ice for 15 mins. Electroporation was done using a Bio-Rad Gene Pulser set to 1250 V, 25 µF and 100 Ω. After electroporation, cells were incubated on ice for 15 mins, then recovered into a total of 500 µl Hayflick media and incubated at 37 °C for 4 h. Then, 125 µl of transformed cells were inoculated into T75 culture flasks containing 20 ml Hayflick and supplemented with 20 µg ml^{-1} chloramphenicol.

The transformed cells were grown to mid-log phase, then isolated. To ensure the recovery of planktonic cells, the media was transferred to a 50 ml Falcon tube. The flask was then scraped into 500 µl Hayflick, which was added to the Falcon tube with the media. The sample was centrifuged at 10000 r.p.m. at 4 °C for 10 mins to pellet the cells. The supernatant

was discarded and the cells resuspended in 500 µl chilled EB. A fresh batch of WT M129 cells was also prepared. The cells were homogenized via 10× passages through a 25-gauge syringe needle. Aliquots of 50 µl of the homogenized cells from both conditions were incubated with a pre-chilled 30 µl EB solution containing the 1 pMole of either the pBSK_p438_Sce_Puro, pBSK_p438_Cre_Puro or pBSK_p438_Puro plasmid, and transformed using the previously described settings. Cells were recovered in the same manner, and 125 µl were inoculated into a T75 flask containing 20 ml Hayflick media supplemented with 3.3 µg ml^{-1} puromycin, and incubated at 37 °C for 5 days.

After the incubation period, cells were isolated via the centrifugation protocol specified above. The pellets were dissolved in 500 µl of fresh Hayflick media, and serial diluted. 10 µl spots of each dilution were then plated on Hayflick agar plates, and incubated at 37 °C for 1 week before being counted.

RESULTS

A large number of papers were omitted from the initial pool of papers, are two main reasons for this. First, all of the papers found with the searches for specific lox sites were also found in the 'Mutant lox site' search, indicating that the remaining papers were more tangentially related to the topic. While almost all papers from the first three searches were relevant, very few papers from the final search were relevant that had not already been accounted for. Second, the 'Mutant lox site' search returned a large number of papers that were not germane to the topic of this review. They generally fit into three further categories: (i) papers involving mutant lox sites that were not lox66 or lox71; (ii) papers describing mutants created by protocols using other lox sites, typically loxP or loxM variants; (iii) papers involving LOX genes, typically from *Pseudomonas aeruginosa*. The classifications of the final 47 papers identified here can be found in Table 2.

In total, we found that 8/47 (17%) of papers surveyed had incorrectly labelled their lox sites. Of these, 7 of the 8 papers [17–23] had simply mislabelled the lox sites, ascribing the lox66 name to the canonical lox71, and vice versa. The remaining papers [24] correctly annotated the lox sites, but implied the directionality of the lox site was in the opposite direction to which it was written.

The strict criteria employed in classifying papers as either correct or incorrect meant that a large number of relevant papers did not fit perfectly into either category. The main reason for this was that they did not explicitly state the sequences of the lox sites that they used. Looking at the figures and methodologies, all of them seem to follow convention and appear to use the lox sites correctly. Furthermore, almost

Table 2. Classification of the 47 relevant papers

Correct papers	Incorrect papers	Unclear papers
References [8, 10, 12, 25, 26, 30–45]	References [17–24]	References [11, 46–62]

every paper references either Albert *et al.*, [8], Leibig *et al.*, [25] or Araki *et al.*, [26] when discussing the addition of lox sites into a construct. All three of these papers show correct annotations for the lox sites that they use.

However, just having the correct reference is not enough to guarantee that the lox site is correct. In Table 1 of Missirlis *et al.*, [21], they show incorrect sequences for lox66 and lox71 while referencing Albert *et al.* [8] and Araki *et al.* [26] for both. Therefore, it seemed reasonable to separate the clear-cut cases from the ambiguous ones, as to create a 'gold standard' of reporting in relation to lox-site nomenclature.

We have shown that there are a large number of errors found within the relevant lox literature. The main cause of these errors potentially stems from the reporting of the original sequences in the paper by Albert *et al.* [8]. In Fig. 3 of their paper, the sequences are shown of the relevant lox sites. Here, the directionality of the lox is facing right to left. However, many of the papers with an error depict the lox site running left to right. In essence, they have switched the directionality of the central spacer region, without also changing which flanking arm contains a deletion.

In the original paper, Albert *et al.* [8] describes the sites as both lox66 and lox71, but also as 'a site with a mutation in the [left element] (LE mutant site) and a site with a mutation in the [right element] (RE mutant site)'. However, the paper does not openly state which is which. In turn the reader is forced to interpret this as either left or right with regard to the directionality of the lox site, or 5′ to 3′ of the given sequence. Therefore, while the diagram of the mutant lox sites is clearly demarcated, the addition of the poorly characterized LE and RE annotations is ambiguous, and may have led to the errors described below. In Fig. 4, we show examples of each of these miss-annotations compare to the original Albert *et al.* [8] descriptions of the lox sites.

It is worth emphasizing at this point that these errors do not affect the scientific validity of the papers they are found in, as the constructs each author describes will create inactive lox sites. It is simply that the annotations the authors attribute the lox sites they use are incorrect. In the papers found with errors in lox-site annotations, they fall into two distinct categories; (i) mis-annotation of the lox-site name, and (ii) mis-annotation of the direction of the lox site.

A good example of the lox sites being simply mis-annotated is shown by Carter and Delneri [17]. In their paper, they refer to the lox sites as left element and right element mutants to try and avoid confusion. Their premise is to flank genes for deletion with mutant lox sites, induce a deletion and create a silent lox72 site, as shown in Fig. 1 of their paper, reproduced below:

Fig. 4. Common examples of lox mis-annotation, showing correct and incorrect examples for each directionality. All mutant regions are underlined. (a) Taken from Albert *et al.* [8]. Shows the original description of the lox66, lox71 and lox72 in relation to the loxP. (b) Sequences taken from Leibig *et al.* [25]. (c) Sequences taken from Weng *et al.* [23]. (d) Sequences taken from Guan *et al.* [20].
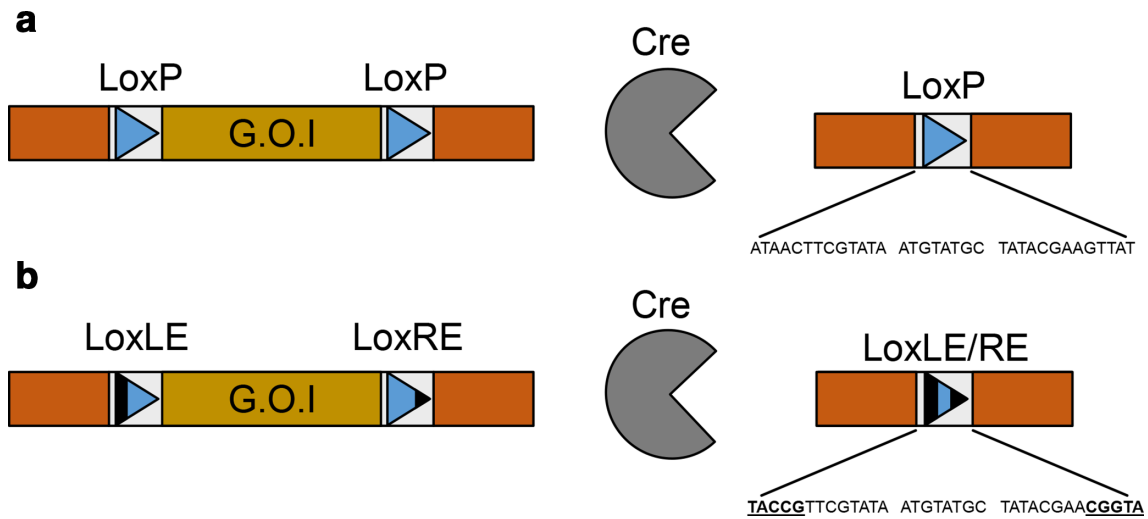
Fig. 5. Modified version of Fig. 1 from Carter and Delneri [17] showing the schematic of their Cre-lox deletion protocol, where their gene of interest (G.O.I) is flanked by different lox sites. (a) The action of the Cre on the DNA causes a deletion between the two loxP sites, leaving a loxP in the genome. (b) The action of the Cre on the DNA leaves an inactive double mutant lox site.

As shown in Fig. 5, the lox sites are inserted running left to right, as indicated by the arrows, with the 5′ or left element of the first lox mutated and the 3′ right element of the second lox mutated. This will give the mutant lox site shown in the figure. However, when looking at the table of oligos provided in Table 1 of the paper, reproduced below, the oligos used to create the lox sites do not match with their role in Fig. 5.

The 'Note' column of Table 3 shows the authors' annotations for the lox sites. However, both are incorrect, as when the lox site is written in a left to right directionality, the left arm mutant is the lox66, not the lox71 [25]. The lox sites are still in the correct orientation, and the correct arms are mutated in each instance. Therefore, while the reaction schematic provided by Carter and Delneri is correct, and will give rise to a mutant lox72 site, the original lox sites are simply mis-labelled.

An example of attributing the wrong direction to the lox site can be found with Weng *et al.* [23]. Fig. 6 is an adaptation

of Fig. 1 from their paper. It shows a self-existing construct flanked by mutant lox sites, with their sequences shown below. The figure shows the lox sites in a left to right orientation, however the sequences of the lox sites shown below it are in a right to left direction. Therefore, the lox66 and lox71 annotations are incorrect. However, this will not affect the validity of the experimental design, as a lox72 will still be formed upon the action of the Cre. The lox66 site described as being in a left to right orientation is in truth a lox71 site in right to left orientation, with the opposite being true for the lox71 site. Simply, the arrows in the figure should be shown in the opposite direction, with the mutant regions of the lox sites still on the 5′ and 3′ arms, respectively.

Lethality conferred by single lox sites

We were interested in examining the potential effects of mis-labelling the lox sites on experimental viability. If one or more lox sites are not correct, or are not internally consistent, then there is the possibility of an active lox site being formed instead of a lox72, likely due to an inversion between the lox sites taking place instead of a deletion. The most damaging scenario would be both the mis-annotation of the lox sites, and their directionality. This occurred within our own laboratory, as a result of two separate people designing plasmids and using different papers as references. The plasmids were intended to produce a lox72 after a deletion had occurred. However, due to the errors in annotation, the resulting constructs were incompatible to create a lox72 and resulted in the formation of a loxP. When this occurred, our attempts at creating a deletion inevitably resulted in cell death, and resulted in months of troubleshooting before its resolution.

This is not encountered in the literature, though this may be via experiments failing and not generating publishable results. As such, we decided to test the effect of a single lox

Table 3. Modification of Table 1 from Carter and Delneri [17]. Annotations to the sequences added for clarity, with spacer unit in bold, and the mutant regions underlined

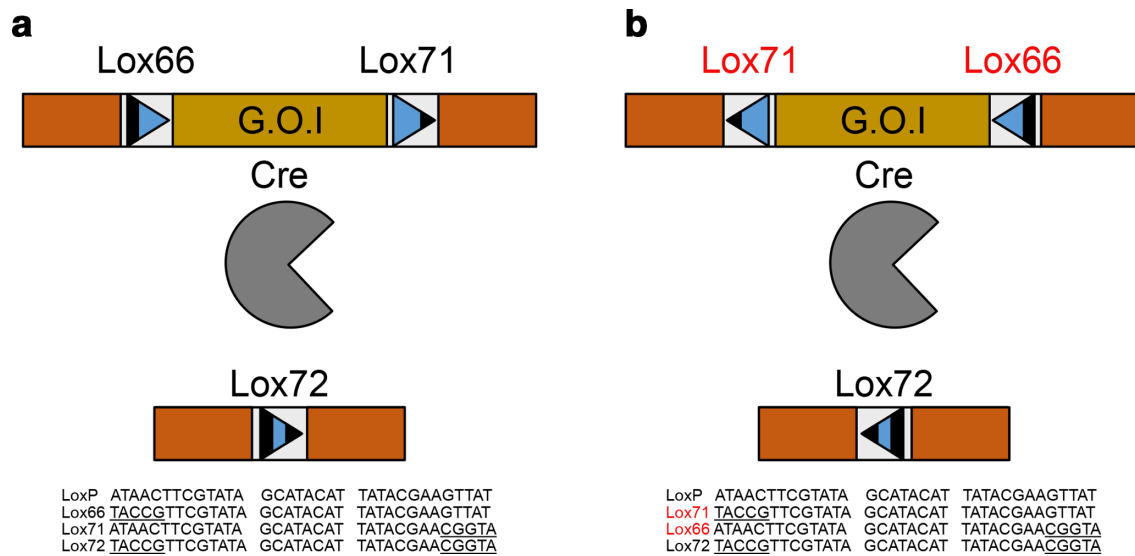| Primer	Sequence (5′ to 3′)	Note
LE_F	GATCCGGGCCC<u>TACCG</u>TTCGTATA **ATGTATGC** TATACGAAGTTATTCTAGAA	Contains lox71 site
LE_R	GATCTTCTAGAATAACTTCGTATA **GCATACAT** TATACGAA<u>CGGTA</u>GGGCCCG	Contains lox71 site
RE_F	CGATACCTGGAAGTCGACATAACTTCGTATA **ATGTATGC** TATACGAA<u>CGGTA</u>GATATCACGTTGCAA	Contains lox66 site
RE_R	CTAGTTGCAACGTGATATC<u>TACCG</u>TTCGTATA **GCATACAT** TATACGAAGTTATGTCGACTTCCAGGTAT	Contains lox66 site

Fig. 6. Simplified experimental schematic, modified from Fig. 1 of Weng *et al.* [23]. (a) Simplification of the original Fig. 1 from Weng *et al.*, [23], showing the directionality and labels attributed to the lox sites used, and their specific sequences. (b) Our modified figure, showing the true annotations and orientations of the construct, with changes in red.

site being present in the genome of a bacteria with the Cre recombinase being expressed, to validate if this annotation error had caused the failure of the experiment.

Using *M. pneumoniae* as a model system, we transformed the cells with a transposon containing a single lox66 site, along with an *I-SceI* meganuclease recognition sequence. This cell line was labelled P0.1. The *I-SceI* meganuclease induces a double-stranded break in the genome, which is lethal to *M. pneumoniae* due to its limited DNA repair mechanisms [27], and provides a good baseline control for lethality. While the usage of the *I-SceI* meganuclease has not been previously demonstrated in *M. pneumoniae*, it has been expressed in other mollicute species, such as *Spiroplasma citri* [28], where it was also used to induce a lethal phenotype.

The *M. pneumoniae* cells were transformed with a transposon containing the *I-SceI* meganuclease recognition site, a lox66 site, and a chloramphenicol resistance gene. This pool of transposon mutants was recovered and then transformed with a suicide vector containing either a Cre recombinase gene, the *I-SceI* meganuclease, or an empty control, alongside the WT cells. A pool of diverse transposon mutants was used instead of a single clone to ensure that the location of the lox site or *I-SceI* site within the genome was not biasing the results.

As shown in Fig. 7, there is a clear lethal phenotype exhibited by the action of the Cre on cell lines with a single lox site present, as after 5 days of incubation with the relevant plasmids. We demonstrate a reduction of 3 orders of magnitude in cell viability measured by c.f.u. (raw data shown in Table S3), which is nearly identical to the lethality induced by the double stranded break enacted by the *I-SceI*. While there is a slight decrease in cell viability between the controls, the WT growing

slightly better than the P0.1 strain, this could be accounted for by the mutant strain having a slightly lower overall fitness due to the presence of the transposon. The *t*-test shows a barely significant difference ($P=0.04$) and is clearly dissimilar to the stark difference between the experimental conditions.

We were confident that individual cells contained only a single transposon, as transformation with 1 pMole of the pMTn_Lox66_Sce_Cm plasmid (3 μg DNA) yields an approximate transformation efficiency of 3×10^{-3}, and successful transformation with two different plasmids lowers the efficiency by an order of magnitude (Fig. S1). It seems likely therefore that the vast majority of cells contained just a single insertion, and thus a single lox or *I-SceI* site.

We therefore demonstrate that both the action of the Cre on a single lox site can cause cell death in *M. pneumoniae*, and also that the *I-SceI* meganuclease can be expressed and active within the system.

DISCUSSION

Due to the fact that at least 17% of the published literature concerning lox66 and lox71 sites contains an annotation error, we felt the need for a more strict naming convention is necessary. While the errors in the published articles do not invalidate them in any way, these inconsistencies in the literature have the potential to severely undermine future experimental designs. If the identity of a lox site is incorrectly determined, then a deletion that aimed to create an inactive lox72 could instead create an active loxP. We have shown that the action of the Cre upon a single active lox site within the genome can have severe consequences for the cell, showing a lethal phenotype in *M. pneumoniae*.

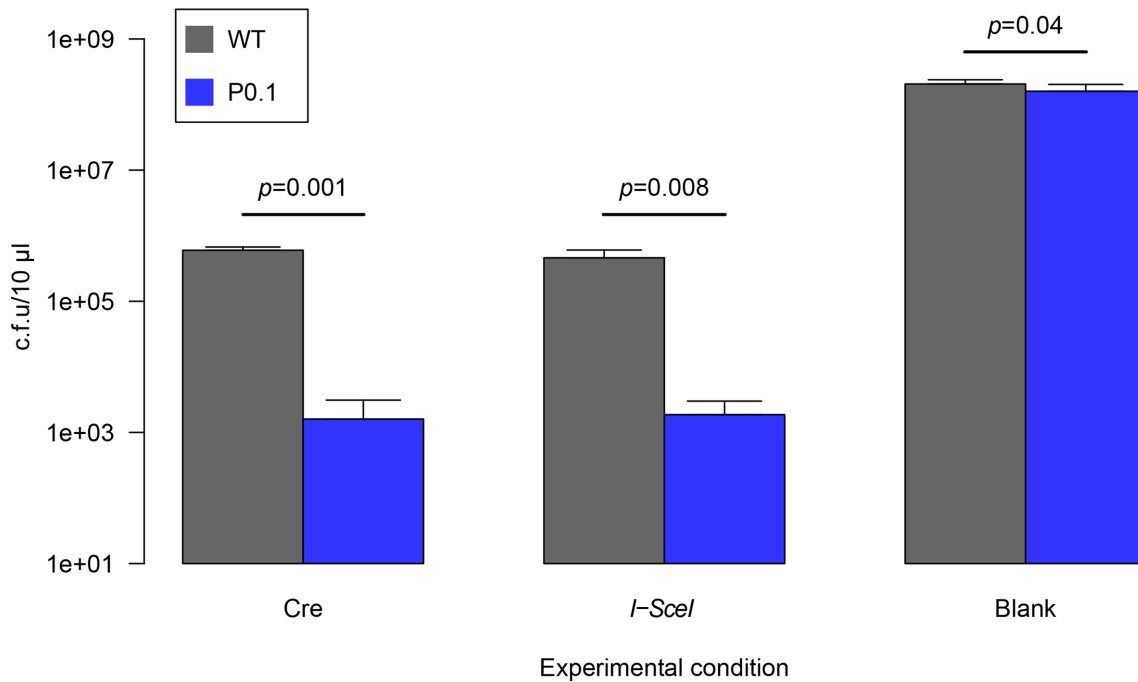**Change in c.f.u after transformation with different plasmids**

Fig. 7. Lethality of a single lox site in *M. pneumoniae* compared to that of the meganuclease *I-SceI*. WT *M. pneumoniae* (grey) and the P0.1 strain (blue), containing a lox66 site and the *I-SceI* recognition site, were transformed with suicide vectors containing a constitutive promoter expressing either the Cre recombinase, the *I-SceI* meganuclease or an empty vector. Bar plots show the changes in the c.f.u. count across the different conditions, along with the significance of the *t*-test used to assay them. Error bars indicate two standard deviations from the mean.

This has the potential to result in the termination of other experiments, especially those where plasmids or methods may be inherited from other researchers, and thus it may not be obvious that sequences for different lox sites have come from separate, incompatible sources. Due to these relatively common errors in annotation, it does not seem unreasonable that other research groups have experienced them. As such, the fact that certain lox resulted in failure may never have made it into the current body of literature, either due to perceived difficulties in publishing negative results or some embodiment of the file-drawer effect. It is our hope that a more standardized naming convention can help in reducing this probability.

Fig. 7 shows that the Cre is highly lethal in *M. pneumoniae* cells that contain a single active lox site, however the cause of this lethality is unknown. We currently have two main hypotheses. The first hypothesis is that the Cre attempts the reaction despite the second lox site not being present. The cleaved leading strand then has no receptor, and thus a nick is formed. If the lagging strand is also cut, then this will form a double-stranded break in the DNA, which is highly lethal in most bacteria [29]. The second hypothesis is that the Cre becomes irreversibly bound to the DNA, as the reaction stalls due to lack of components. This may leave the Cre bound to the DNA during the strand cleavage reactions, and thus not only creates nicks and potential double-strand breaks, but also becomes a hindrance to DNA transcription, translation and replication machinery.

While our experiments showed a severe lethal phenotype in bacteria with a known deficiency in their DNA damage repair mechanisms, the phenomena could also be relevant to other species. Regardless of what the underlying cause of the lethality is, if the Cre is constitutively expressed then the damage it causes may be enough to overwhelm the native repair mechanisms, or the constant repair needed may be a sufficient metabolic burden to result in a severe loss of cellular fitness. Similarly, if the Cre interferes with DNA transcription or replication through its binding to the DNA, then it could cause strong effects in bacteria with more robust DNA repair mechanisms as well.

Despite this, it is apparent that this action also has beneficial applications in genome engineering, as either a kill-switch mechanism or selection against cells that have undergone unwanted DNA inversions. This was demonstrated by Shaw *et al.* (in preparation) as a highly effective counter-selection system to remove unwanted inversions when lox sites were added randomly to the genome of *M. pneumoniae*, leaving

only cells that had undergone a deletion and produced an inactive lox72 site, as cells that contained inversions, which produced active lox sites were killed.

CONCLUSIONS

Of the 21 articles that unambiguously showed the correct lox-site annotations, 18/21 used the same lox orientation as the original paper by Albert *et al.* [8], with the lox directionality reading from right to left [8, 12, 26, 30–45]. In contrast, five of the eight papers with mis-labelled lox sites portray them from left to right [17, 20–22, 24]. It seems apparent that there is potential for the mis-annotation of these sequences, and that this could stem from incorrectly assuming either lox-site directionality, or by accidentally attributing the wrong name to the correct site.

While the cases identified here are benign, with mis-annotations not affecting the outcomes of the experiments, the prevalence of these errors is worrying. Attempts at replication, or researches using these maps or constructs as the basis for their own experiments may inadvertently create incompatible designs of vectors, and thus cause delays and errors in future experiments. We have demonstrated that the action of the Cre on a single lox site can have a lethal effect, and therefore the authors believe that special care should be taken to properly attribute the correct names to lox sites.

As such, it seems prudent to recommend that researchers continue using the annotations provided by Albert *et al.* [8], repeated below, as this is the form taken by the majority of the correct cases, and unused by the majority of the incorrect ones. We also recommend the avoidance of the left or right element notation, as it is too ambiguous and dependent on orientation. If the lox site is written from right to left, then the LE mutant would be the lox71, however if the lox site is written left to right, then the LE becomes lox66. Therefore, in the aim of standardization, we recommend following the original notation and designing lox sites from right to left whenever possible:

Lox66: 5′ - ATAACTTCGTATA GCATACAT TATAC-GAAcggta - 3′

←

Lox71: 5′ - taccgTTCGTATA GCATACAT TATACGAA-GTTAT - 3′

By adopting a more standardized nomenclature, hopefully the internal reproducibility and accuracy of the reporting of lox sites, along with errors or confusion caused by the usage of incorrect sites, can be minimized.

Author contributions
D.S. conceptualization, data curation, formal analysis, investigation, methodology, validation, visualization, writing – original draft, writing – review and editing. L.S. conceptualization, funding acquisition, resources, supervision, writing – review and editing. M.L–S. conceptualization, supervision, writing – review and editing.

References

1. Sternberg N, Hamilton D. Bacteriophage P1 site-specific recombination. I. recombination between loxP sites. *J Mol Biol* 1981;150:467–486.

2. Sternberg N, Hamilton D, Austin S, Yarmolinsky M, Hoess R. Site-specific recombination and its role in the life cycle of bacteriophage P1. *Cold Spring Harb Symp Quant Biol* 1981;45 Pt 1:297–309.

3. Ghosh K, Van Duyne GD. Cre-loxP biochemistry. *Methods* 2002;28:374–383.

4. Yarmolinsky M, Hoess R. The legacy of NAT Sternberg: the genesis of cre-lox technology. *Annu Rev Virol* 2015;2:25–40.

5. Van Duyne GD. Cre recombinase. *Microbiol Spectr* 2015 Feb;3: MDNA3-0014–3-2014.

6. Song AJ, Palmiter RD. Detecting and avoiding problems when using the cre-lox system. *Trends Genet* 2018;34:333–340.

7. Sauer B. Cre/Lox: one more step in the taming of the genome. *Endocrine* 2002;19:221–228.

8. Albert H, Dale EC, Lee E, Ow DW, . Site-specific integration of DNA into wild-type and mutant LOX sites placed in the plant genome. *Plant J* 1995;7:649–659.

9. Gelato KA, Martin SS, Wong S, Baldwin EP. Multiple levels of affinity-dependent DNA discrimination in Cre-LoxP recombination. *Biochemistry* 2006;45:12216–12226.

10. Lambert JM, Bongers RS, Kleerebezem M. Cre-lox-based system for multiple gene deletions and selectable-marker removal in *Lactobacillus plantarum*. *Appl Environ Microbiol* 2007;73:1126–1135.

11. Zhang S, Ban A, Ebara N, Mizutani O, Tanaka M *et al*. Self-excising Cre/mutant lox marker recycling system for multiple gene integrations and consecutive gene deletions in *Aspergillus oryzae*. *J Biosci Bioeng* 2017;123:403–411.

12. Suzuki N, Nonaka H, Tsuge Y, Inui M, Yukawa H. New multiple-deletion method for the *Corynebacterium glutamicum* genome, using a mutant lox sequence. *Appl Environ Microbiol* 2005;71:8472–8480.

13. Hayflick L. Tissue cultures and mycoplasmas. *Tex Rep Biol Med* 1965;23:285.

14. Yus E, Maier T, Michalodimitrakis K, van Noort V, Yamada T *et al*. Impact of genome reduction on bacterial metabolism and its regulation. *Science* 2009;326:1263–1268.

15. Gibson DG, Young L, Chuang R-Y, Venter JC, Hutchison CA *et al*. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Methods* 2009;6:343–345.

16. Hedreyda CT, Lee KK, Krause DC. Transformation of *Mycoplasma pneumoniae* with Tn4001 by electroporation. *Plasmid* 1993;30:170–175.

17. Carter Z, Delneri D. New generation of loxP-mutated deletion cassettes for the genetic manipulation of yeast natural isolates. *Yeast* 2010;27:765–775.

18. Cerisy T, Rostain W, Chhun A, Boutard M, Salanoubat M *et al*. A Targetron-Recombinase system for large-scale genome engineering of clostridia. *mSphere* 2019 Dec 11;4

19. Erler A, Maresca M, Fu J, Stewart AF. Recombineering reagents for improved inducible expression and selection marker re-use in *Schizosaccharomyces pombe*. *Yeast* 2006;23:813–823.

20. Guan Z-B, Wang K-Q, Shui Y, Liao X-R. Establishment of a markerless multiple-gene deletion method based on Cre/loxP mutant system for *Bacillus pumilus*. *J Basic Microbiol* 2017;57:1065–1068.

21. 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;7:73.

22. Parrish M, Unruh J, Krumlauf R. Bac modification through serial or simultaneous use of Cre/lox technology. *J Biomed Biotechnol* 2011;2011:1–12

23. Weng L, Biswas I, Morrison DA. A self-deleting Cre-lox-ermAM cassette, Cheshire, for marker-less gene deletion in *Streptococcus pneumoniae*. *J Microbiol Methods* 2009;79:353–357.

24. Chatterjee PK, Shakes LA, Stennett N, Richardson VL, Malcolm TL *et al.* Replacing the wild type loxP site in BACs from the public domain with lox66 using a lox66 transposon. *BMC Res Notes* 2010;3:38.

25. Leibig M, Krismer B, Kolb M, Friede A, Götz F *et al.* Marker removal in staphylococci via Cre recombinase and different LOX sites. *Appl Environ Microbiol* 2008;74:1316–1323.

26. Araki K, Araki M, Yamamura K. Targeted integration of DNA using mutant LOX sites in embryonic stem cells. *Nucleic Acids Res* 1997;25:868–872.

27. Razin S, Yogev D, Naot Y. Molecular biology and pathogenicity of mycoplasmas. *Microbiol Mol Biol Rev* 1998;62:1094–1156.

28. Breton M, Duret S, Béven L, Dubrana M-P, Renaudin J. I-SceI-mediated plasmid deletion and intra-molecular recombination in *Spiroplasma citri*. *J Microbiol Methods* 2011;84:216–222.

29. Ayora S, Carrasco B, Cárdenas PP, César CE, Cañas C *et al.* Double-strand break repair in bacteria: a view from *Bacillus subtilis*. *FEMS Microbiol Rev* 2011;35:1055–1081.

30. Araki K, Araki M, Yamamura K-ichi. Site-Directed integration of the Cre gene mediated by Cre recombinase using a combination of mutant LOX sites. *Nucleic Acids Res* 2002;30:e103–.

31. Araki K, Okada Y, Araki M, Yamamura K-ichi. Comparative analysis of right element mutant LOX sites on recombination efficiency in embryonic stem cells. *BMC Biotechnol* 2010;10:29.

32. Bertram R, Kolb M, Hillen W. In vivo activation of tetracycline repressor by Cre/lox-mediated gene assembly. *J Mol Microbiol Biotechnol* 2009;17:136–145.

33. Jiang Y, Gao X, Xu K, Wang J, Huang H *et al.* A Novel Cre Recombinase-Mediated *In Vivo* Minicircle DNA (CRIM) Vaccine Provides Partial Protection against Newcastle Disease Virus. *Appl Environ Microbiol* 2019;85

34. Koussis K, Withers-Martinez C, Baker DA, Blackman MJ. Simultaneous multiple allelic replacement in the malaria parasite enables dissection of PKG function. *Life Sci Alliance* 2020;3:e201900626 [Epub ahead of print 16 03 2020].

35. Langer SJ, Ghafoori AP, Byrd M, Leinwand L. A genetic screen identifies novel non-compatible loxP sites. *Nucleic Acids Res* 2002;30:3067–3077.

36. Liu W-yi, Wang Y, Qin Y, Wang Y-ping, Zhu Z-yan, Wang Y, Zhu Z. Site-Directed gene integration in transgenic zebrafish mediated by Cre recombinase using a combination of mutant LOX sites. *Mar Biotechnol* 2007;9:420–428.

37. Oberdoerffer P, Otipoby KL, Maruyama M, Rajewsky K. Unidirectional Cre-mediated genetic inversion in mice using the mutant loxP pair lox66/lox71. *Nucleic Acids Res* 2003;31:e140–.

38. Pan R, Zhang J, Shen W-L, Tao Z-Q, Li S-P, *et al.* Sequential deletion of *Pichia pastoris* genes by a self-excisable cassette. *FEMS Yeast Res* 2011;11:292–298.

39. Petersen KV, Martinussen J, Jensen PR, Solem C. Repetitive, marker-free, site-specific integration as a novel tool for multiple chromosomal integration of DNA. *Appl Environ Microbiol* 2013;79:3563–3569.

40. Suzuki N, Inui M, Yukawa H. Site-directed integration system using a combination of mutant lox sites for *Corynebacterium glutamicum*. *Appl Microbiol Biotechnol* 2007;77:871–878.

41. Tomimoto K, Yamakawa M, Tanaka H. Construction of a long hairpin RNA expression library using Cre recombinase. *J Biotechnol* 2012;160:129–139.

42. Wang Y, Weng J, Waseem R, Yin X, Zhang R *et al. Bacillus subtilis* genome editing using ssDNA with short homology regions. *Nucleic Acids Res* 2012;40:e91.

43. Wu L, Wu H, Chen L, Lin L, Borriss R *et al.* Bacilysin overproduction in *Bacillus amyloliquefaciens* FZB42 markerless derivative strains FZBREP and FZBSPA enhances antibacterial activity. *Appl Microbiol Biotechnol* 2015;99:4255–4263.

44. Yan X, Yu H-J, Hong Q, Li S-P, H-J Y, S-P L. Cre/lox system and PCR-based genome engineering in *Bacillus subtilis*. *Appl Environ Microbiol* 2008;74:5556–5562.

45. Zhang Z, Lutz B. Cre recombinase-mediated inversion using lox66 and lox71: method to introduce conditional point mutations into the CREB-binding protein. *Nucleic Acids Res* 2002;30:e90–.

46. Xin Y, Guo T, Mu Y, Kong J. A Single-Plasmid Genome Editing System for Metabolic Engineering of *Lactobacillus casei*. *Front Microbiol* 2018;9:3024.

47. Kimura Y, Ding B, Imai N, Nolan DJ, Butler JM *et al.* c-Kit-Mediated functional positioning of stem cells to their niches is essential for maintenance and regeneration of adult hematopoiesis. *PLoS One* 2011 Oct 26;6:e26918

48. Taniwaki T, Haruna K, Nakamura H, Sekimoto T, Oike Y *et al.* Characterization of an exchangeable gene trap using pU-17 carrying a stop codon-beta Geo cassette. *Dev Growth Differ* 2005;47:163–172.

49. Döhlemann J, Brennecke M, Becker A. Cloning-free genome engineering in *Sinorhizobium meliloti* advances applications of Cre/loxP site-specific recombination. *J Biotechnol* 2016;233:160–170.

50. Xin Y, Guo T, Mu Y, Kong J. Coupling the recombineering to Cre-lox system enables simplified large-scale genome deletion in *Lactobacillus casei*. *Microb Cell Fact* 2018;17:21.

51. Berzin V, Kiriukhin M, Tyurin M. Cre-lox66/lox71-based elimination of phosphotransacetylase or acetaldehyde dehydrogenase shifted carbon flux in acetogen rendering selective overproduction of ethanol or acetate. *Appl Biochem Biotechnol* 2012;168:1384–1393.

52. Togawa Y, Nunoshiba T, Hiratsu K. Cre/lox-based multiple markerless gene disruption in the genome of the extreme thermophile *Thermus thermophilus*. *Mol Genet Genomics* 2018;293:277–291.

53. Kim JM, Lee KH, Lee SY. Development of a markerless gene knockout system for *Mannheimia succiniciproducens* using a temperature-sensitive plasmid. *FEMS Microbiol Lett* 2008;278:78–85.

54. Li Z, Zhao G, Shen J, Araki K, Haruna K *et al.* Enhanced expression of human cDNA by phosphoglycerate kinase promoter-puromycin cassette in the mouse transthyretin locus. *Transgenic Res* 2011;20:191–200.

55. Kiriukhin M, Tyurin M. Expression of amplified synthetic ethanol pathway integrated using Tn7-tool and powered at the expense of eliminated PTA, ACK, Spo0A and spo0J during continuous syngas or CO_2/H_2 blend fermentation. *J Appl Microbiol* 2013;114:1033–1045.

56. Zhao G, Li Z, Araki K, Haruna K, Yamaguchi K *et al.* Inconsistency between hepatic expression and serum concentration of transthyretin in mice humanized at the transthyretin locus. *Genes Cells* 2008;13:1257–1268.

57. Liu X-B, Liu M, Tao X-Y, Zhang Z-X, Wang F-Q *et al.* Metabolic engineering of *Pichia pastoris* for the production of dammarenediol-II. *J Biotechnol* 2015;216:47–55.

58. Kiriukhin M, Tyurin M. Mevalonate production by engineered acetogen biocatalyst during continuous fermentation of syngas or CO_2/H_2 blend. *Bioprocess Biosyst Eng* 2014;37:245–260.

59. Araki K, Araki M, Yamamura K-ichi. Negative selection with the diphtheria toxin a fragment gene improves frequency of Cre-mediated cassette exchange in ES cells. *J Biochem* 2006;140:793–798.

60. Tyurin M, Kiriukhin M. Selective methanol or formate production during continuous CO_2 fermentation by the acetogen biocatalysts engineered via integration of synthetic pathways using Tn7-tool. *World J Microbiol Biotechnol* 2013;29:1611–1623.

61. **Berzin V**, **Tyurin M**, **Kiriukhin M**. Selective n-butanol production by *Clostridium* sp. MTButOH1365 during continuous synthesis gas fermentation due to expression of synthetic thiolase, 3-hydroxy butyryl-CoA dehydrogenase, crotonase, butyryl-CoA dehydrogenase, butyraldehyde dehydrogenase, and NAD-dependent butanol dehydrogenase. *Appl Biochem Biotechnol* 2013;169:950–959.

62. **Tyurin M**, **Kiriukhin M**. Synthetic 2,3-butanediol pathway integrated using Tn7-tool and powered via elimination of sporulation and acetate production in acetogen biocatalyst. *Appl Biochem Biotechnol* 2013;170:1503–1524.

63. **Burgos R**, **Totten PA**. Characterization of the operon encoding the Holliday junction helicase RuvAB from *Mycoplasma genitalium* and its role in mgpB and mgpC gene variation. *J Bacteriol* 2014;196:1608–1618.

Edited by: E. Denham and S. Purton