

RESEARCH ARTICLE

Open Access

# Simplified plasmid cloning with a universal MCS design and bacterial in vivo assembly



Fan Chen\*, Yi-ya Li, Yan-li Yu, Jie Dai, Jin-ling Huang and Jie Lin

## Abstract

**Background:** The ability to clone DNA sequences quickly and precisely into plasmids is essential for molecular biology studies. The recent development of seamless cloning technologies has made significant improvements in plasmid construction, but simple and reliable tools are always desirable for time- and labor-saving purposes.

**Results:** We developed and standardized a plasmid cloning protocol based on a universal MCS (Multiple Cloning Site) design and bacterial in vivo assembly. With this method, the vector is linearized first by PCR (Polymerase Chain Reaction) or restriction digestion. Then a small amount (10 ~ 20 ng) of this linear vector can be mixed with a PCR-amplified insert (5× molar ratio against vector) and transformed directly into competent *E. coli* cells to obtain the desired clones through in vivo assembly. Since we used a 36-bp universal MCS as the homologous linker, any PCR-amplified insert with ~ 15 bp compatible termini can be cloned into the vector with high fidelity and efficiency. Thus, the need for redesigning insert-amplifying primers according to various vector sequences and the following PCR procedures was eliminated.

**Conclusions:** Our protocol significantly reduced hands-on time for preparing transformation reactions, had excellent reliability, and was confirmed to be a rapid and versatile plasmid cloning technique. The protocol contains mostly mixing steps, making it an extremely automation-friendly and promising tool in modern biology studies.

**Keywords:** Plasmid cloning, Universal MCS, Bacterial in vivo assembly, DNA assembly, Homologous sequence

Plasmid cloning is one of the most commonly used techniques in molecular biology research. It plays a crucial role in studying the structure, function, and evolution of genes [1, 2] while serving as an essential tool in genetic, protein, and metabolic engineering [3, 4]. However, the traditional digestion-ligation method is often limited, as both vector and target fragments must have compatible cleavage sites. Moreover, the whole procedure can be time-consuming and labor-intensive [5].

To overcome the difficulties mentioned above, new alternatives have emerged in the past two decades. These include PCR-mediated IVA (In Vivo Assembly) cloning [6], fast cloning [7]; LIC (Ligation-Independent Cloning) such as ELIC (Exonuclease and Ligation-Independent

Cloning) [8], SLIC (Sequence and Ligation-Independent Cloning) [9], HAC (Homologous Alignment Cloning) [10], One-step SLIC [11] and In-Fusion cloning [12]; recombination-based Gateway Cloning [13], SLICE (Seamless Ligation Cloning Extract) [14, 15], and multienzyme-based Gibson Assembly [16], Golden Gate Assembly [17]. Each of these methods has advantages and limitations (more information is listed in Supplementary Table S1). For example, PCR-mediated and LIC methods remove a few enzymatic treatment steps. Nevertheless, multiple rounds of PCR amplification are required, which might result in random mutations in the final products. Recombination- and multienzyme-based methods are often used in commercial cloning kits. They are methods with excellent convenience, but the cost of applying these kits is also much higher. Furthermore, most cloning methods are not automation friendly

\* Correspondence: [cf@bio.mnnu.edu.cn](mailto:cf@bio.mnnu.edu.cn)

School of Biological Science and Biotechnology, Minnan Normal University, Zhangzhou 363000, P.R. China



© The Author(s). 2021 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

because multiple pre-transformation steps or different primer designs are necessary for various vectors and inserts.

The *recA*-independent recombination pathway in *E. coli* (Fig. 1) has been reported for more than two decades [18, 19]. Although the mechanism of such recombination remains elusive, it was recently considered a useful tool for simplifying molecular cloning. When applying such methods, terminal sequences of linearized vectors were always chosen as the homologous linker required by *in vivo* assembly. Thus, as the vector sequence changes, the primers used to amplify the insert will have to be redesigned and synthesized, resulting in additional costs and labor.

In this study, inspired by previous reports [6, 8, 20, 21], we developed a novel cloning strategy based on a UMCS (universal MCS) design and bacterial *in vivo* assembly. The method uses 10 ~ 20 ng of linearized vector and a 5× molar ratio of insert fragment for successful subcloning with high efficiency and fidelity. The whole transformation process can be performed within 1.5 h when carried out using classic chemically competent cell strains. The main concept of our strategy is to design and synthesize a universal MCS that can be used to replace the existing MCS of each vector. When the UMCS serves as the homologous linker, the primers to amplify the vector are determined. Only the template binding section of primers for insert amplification needs to be changed. Due to the consistency of homologous linker sequences, the same amplified DNA fragments can be cloned into any modified vector with the UMCS without extra PCR steps. This makes the method simple, robust, highly versatile, and further meets the requirements of a high-throughput automated operation [22].

## Results

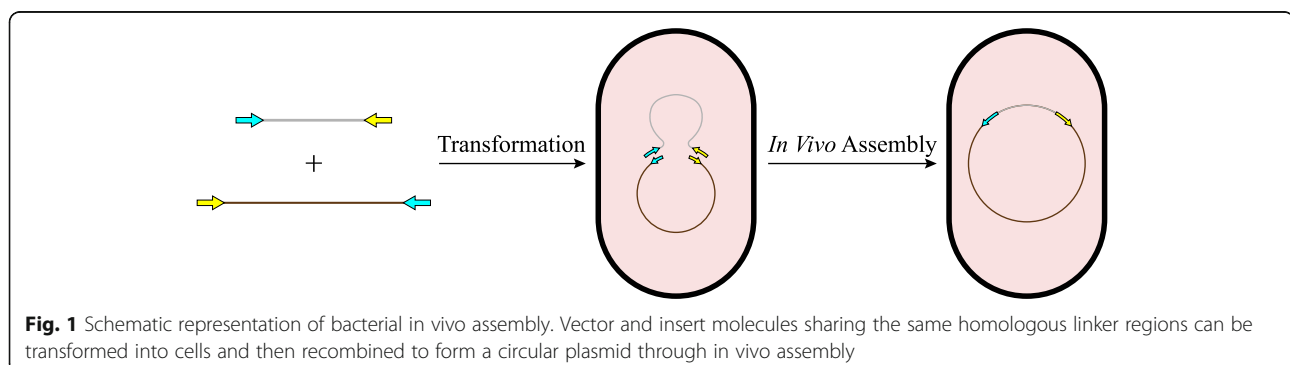
### Universal MCS and its associated primer design

The design of the UMCS for universal subcloning is shown in Fig. 2a. The main features of this novel MCS design are as follows: (1) The length of UMCS is only

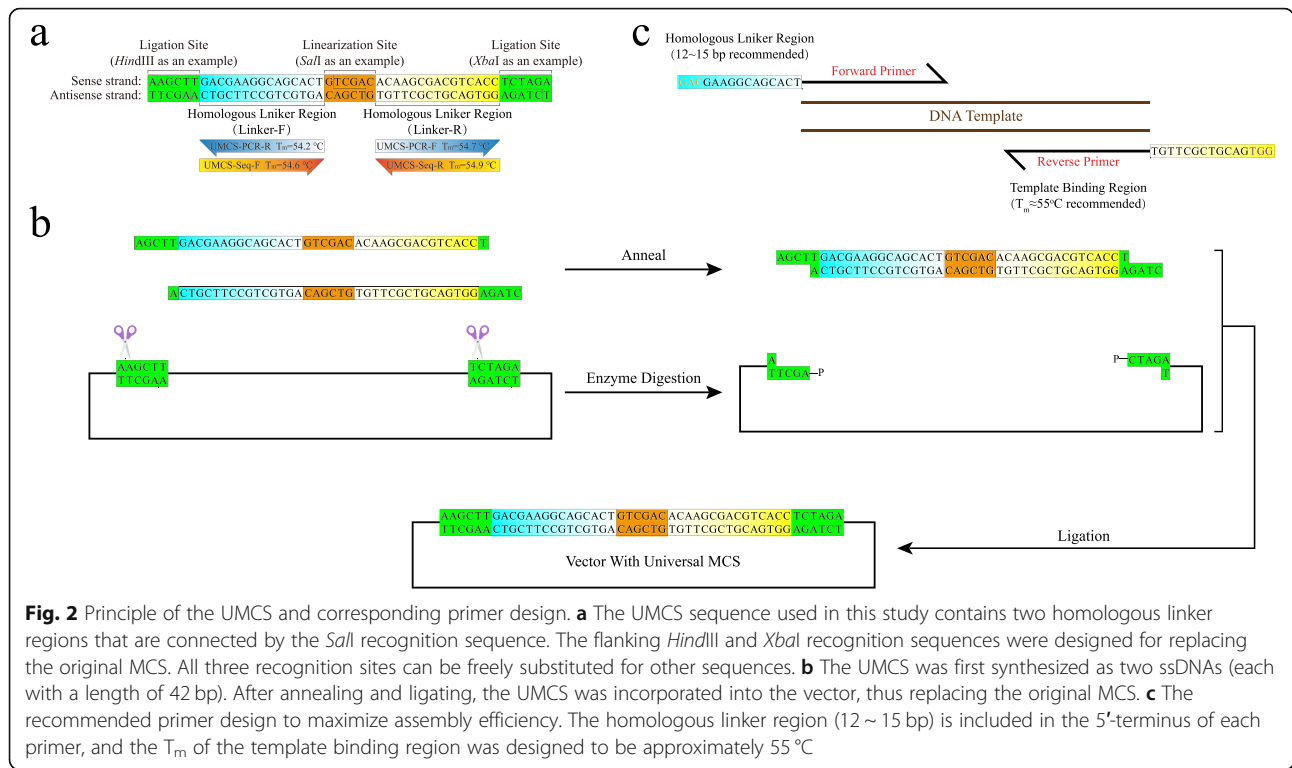
48 bp, which is shorter than many of the existing MCSs. The actual sequence to be synthesized is 42 bp for both the forward and reverse directions (Fig. 2b). By annealing and ligating to a predigested vector molecule, the plasmid with the UMCS was reconstituted. (2) The restriction enzyme recognition sites at both termini are needed only for UMCS integration. They can be replaced freely according to the original MCS compatibility of the vector. (3) The restriction enzyme recognition site in the middle of the UMCS is used only for vector linearization, which can be substituted with any other sequences for digestion if they are compatible with the vector. (4) The two homologous linker regions were carefully optimized to have close  $T_m$  and the same GC content, which are supposed to facilitate PCR. Hairpin and dimerization of these regions were also minimized through optimization studies. (5) Primers used for the amplification of the insert can be easily designed by introducing a homologous linker at their 5'-terminus, while the  $T_m$  of the template binding region is calculated to be approximately 55 °C (Fig. 2c). (6) Theoretically, because of MCS consistency, any vector with a UMCS sequence can be amplified by the "UMCS-PCR-F + UMCS-PCR-R" primer pair, and any insert cloned into the UMCS can be amplified by the "UMCS-Seq-F + UMCS-Seq-R" primer pair. (7) Both linker regions do not contain rare codons [23, 24]. If the linker sequence is to be expressed, the codons contained in the UMCS are unlikely to affect the target protein expression. This is crucial for vectors that express N-terminal tags.

### Construction of pUC19-UMCS-EGFP

To demonstrate proof of principle, we first constructed two types of pUC19-UMCS-EGFP vectors (Fig. 3) as tools for studying the transformation efficiency. These vectors are modified versions of the pUC19 because the *lacZ* sequence was first replaced by UMCS-*XhoI* sequences with different linearization sites (*Sall* or *EcoRV*, as shown in Fig. 3a and 3b), followed by the insertion of EGFP (Enhanced Green Fluorescent Protein)



**Fig. 1** Schematic representation of bacterial *in vivo* assembly. Vector and insert molecules sharing the same homologous linker regions can be transformed into cells and then recombined to form a circular plasmid through *in vivo* assembly

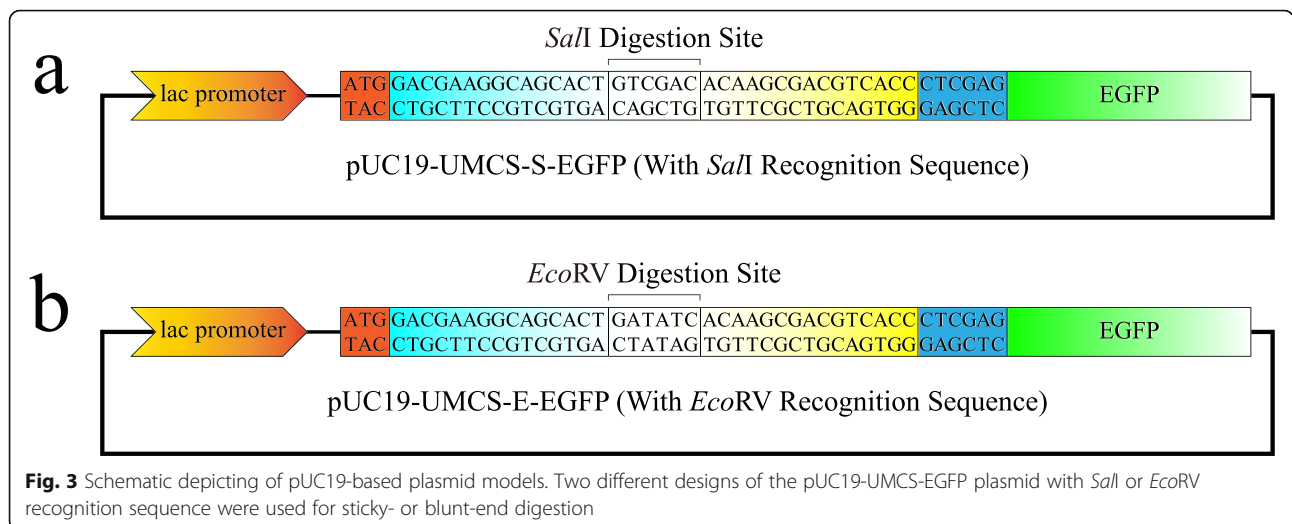


**Fig. 2** Principle of the UMCS and corresponding primer design. **a** The UMCS sequence used in this study contains two homologous linker regions that are connected by the *Sal*I recognition sequence. The flanking *Hind*III and *Xba*I recognition sequences were designed for replacing the original MCS. All three recognition sites can be freely substituted for other sequences. **b** The UMCS was first synthesized as two ssDNAs (each with a length of 42 bp). After annealing and ligating, the UMCS was incorporated into the vector, thus replacing the original MCS. **c** The recommended primer design to maximize assembly efficiency. The homologous linker region (12 ~ 15 bp) is included in the 5'-terminus of each primer, and the  $T_m$  of the template binding region was designed to be approximately 55 °C

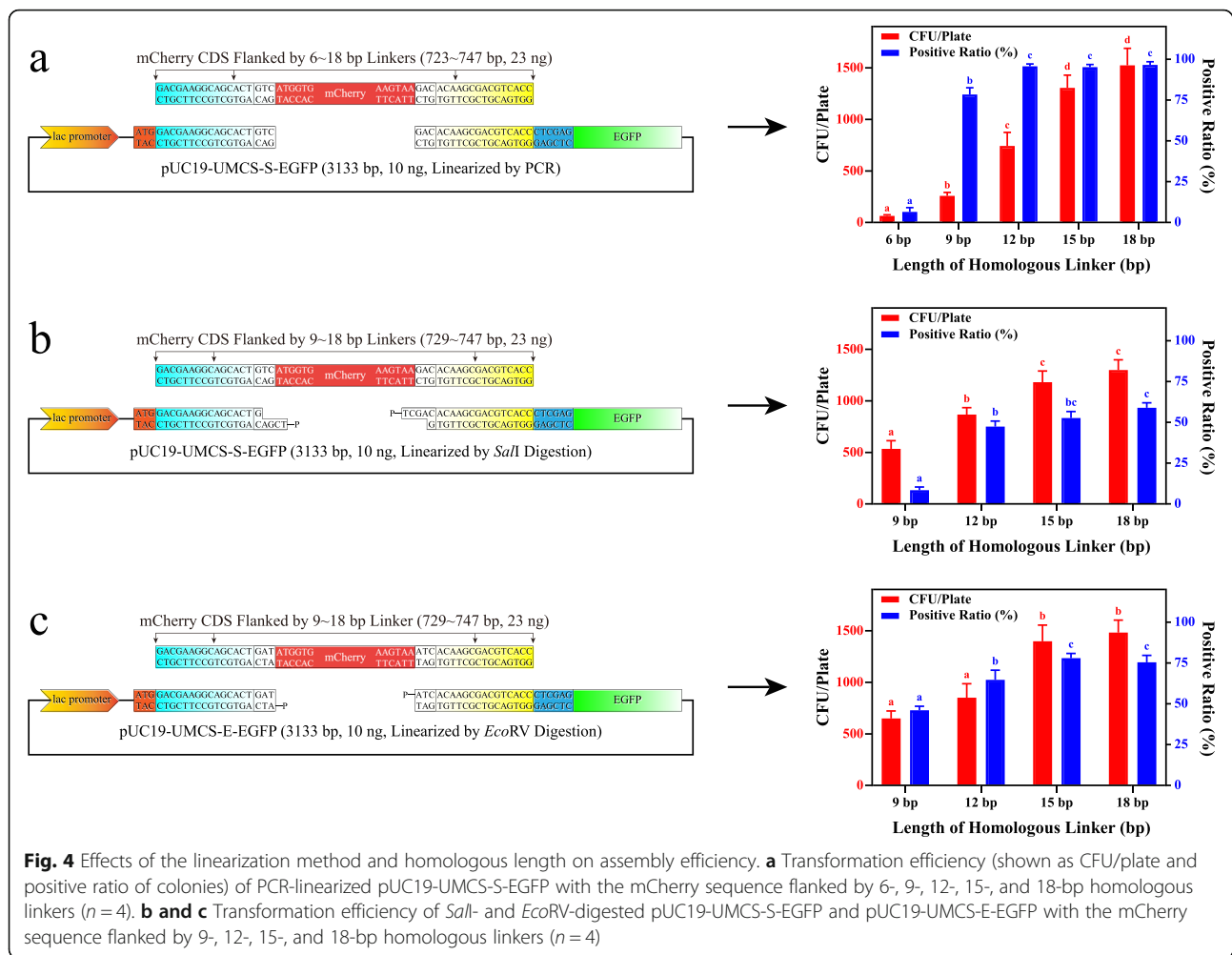
through *Xho*I-based digestion and ligation. Among them, *Sal*I represented the linearized product with sticky ends, while *Eco*RV represented blunt ends. Together with vectors linearized by PCR, the effect of different linearization methods on transformation can be clarified. During in vivo assembly, if the mCherry CDS is successfully cloned into the UMCS, the colony will appear orange. Otherwise, it will be green or colorless (through nonspecific assembly). By calculating orange colonies against the total number of colonies, the positive ratio of transformation can be determined.

### Effects of the linearization method and homologous linker length

Next, we evaluated the effects of different linearization methods and the length of the homologous linker in terms of the number of transformants and the corresponding positive ratio. As shown in Fig. 4a, when PCR was used as a linearization method, the transformation efficiency showed a positive correlation with the length of the linker sequence. When a 6-bp linker was used, less than 100 CFU/plate and a positive ratio below 10% were confirmed. In contrast, the efficiency of



**Fig. 3** Schematic depicting of pUC19-based plasmid models. Two different designs of the pUC19-UMCS-EGFP plasmid with *Sal*I or *Eco*RV recognition sequence were used for sticky- or blunt-end digestion



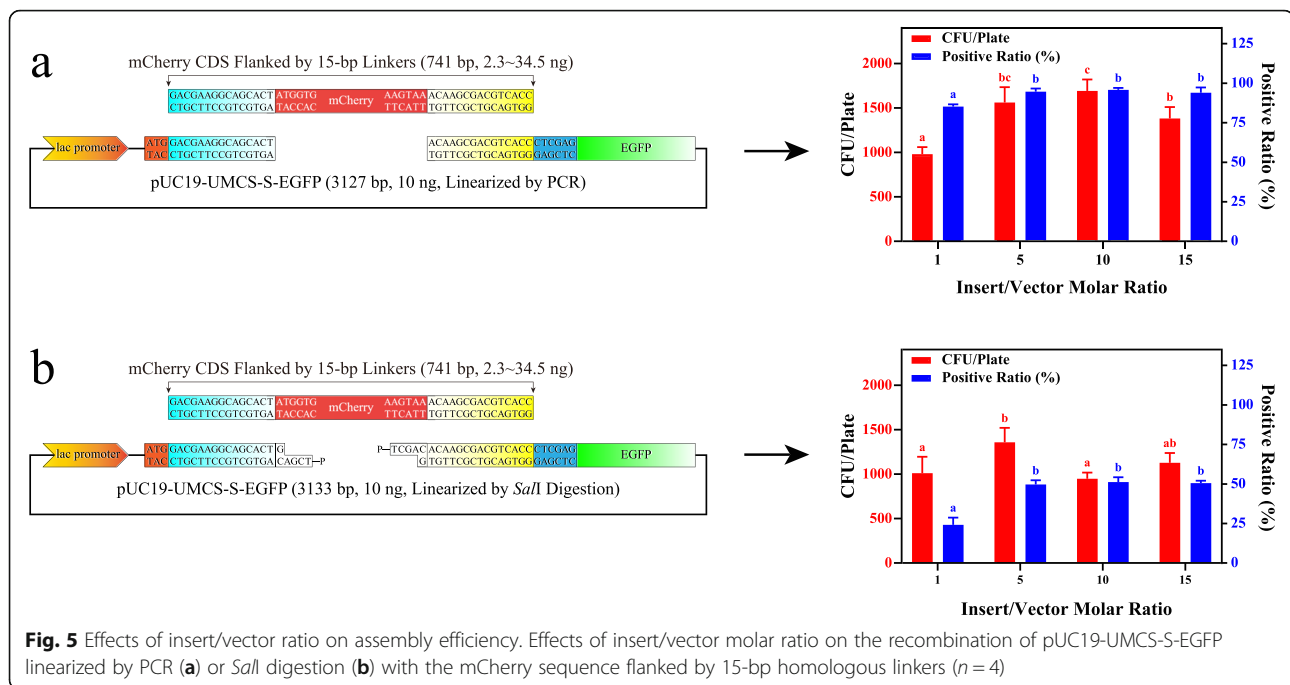
transformation increased to  $259 \pm 34$ ,  $742 \pm 132$ ,  $1306 \pm 123$ , and  $1525 \pm 165$  CFU/plate, and the positive ratio increased dramatically to  $78 \pm 4.2\%$ ,  $96 \pm 1.5\%$ ,  $95 \pm 1.7\%$ , and  $97 \pm 1.9\%$  when 9-, 12-, 15-, and 18-bp linkers were tested, respectively. It is evident that a 6-bp linker is not enough for bacterial in vivo assembly, and a length longer than 9 bp is necessary when performing such experiments. Moreover, there were no significant differences between the 15- and 18-bp linker. Considering both the colony number and positive ratio, we concluded that the length of the homologous linker for the PCR method was 12 ~ 15 bp. On the other hand, when linearization was carried out by restriction digestion, it was shown (Fig. 4b and 4c) that the digested vectors with blunt ends (*EcoRV*) yielded much more positive clones than sticky ends at any of the linker lengths we tested. According to the data collected, when a restriction enzyme was used for linearization, a 12 ~ 15-bp linker also ensured enough transformants and positive clones to screen. In these cases, further increase the linker length is not necessary.

#### Effects of insert/vector ratio on transformation

The molar ratio of insert/vector is considered a critical factor for high-efficiency transformation [25]. Thus, we further investigated the influence of PCR, *Sall* and *EcoRV* digestion on transformation efficiency at multiple insert/vector levels. The yields of transformants and the positive ratio at the linker length of 15 bp were determined (Fig. 5 and Supplementary Figure S1) with molar ratios of 1, 5, 10, and 15. According to the data shown, regardless of the linearization method used, the optimal molar ratio remained at 5. A further increase in insert usage will not improve the results, while a decreased insert quantity shows a negative effect on transformation efficiency. Therefore, we used a molar ratio of 5 in all subsequent studies, which is consistent with the finding of Kostylev et al. [26]

#### Effects of the digestive linearization method on the fidelity of assembly

We were interested in the assembly fidelity of the digestive linearization method mainly based on three reasons:

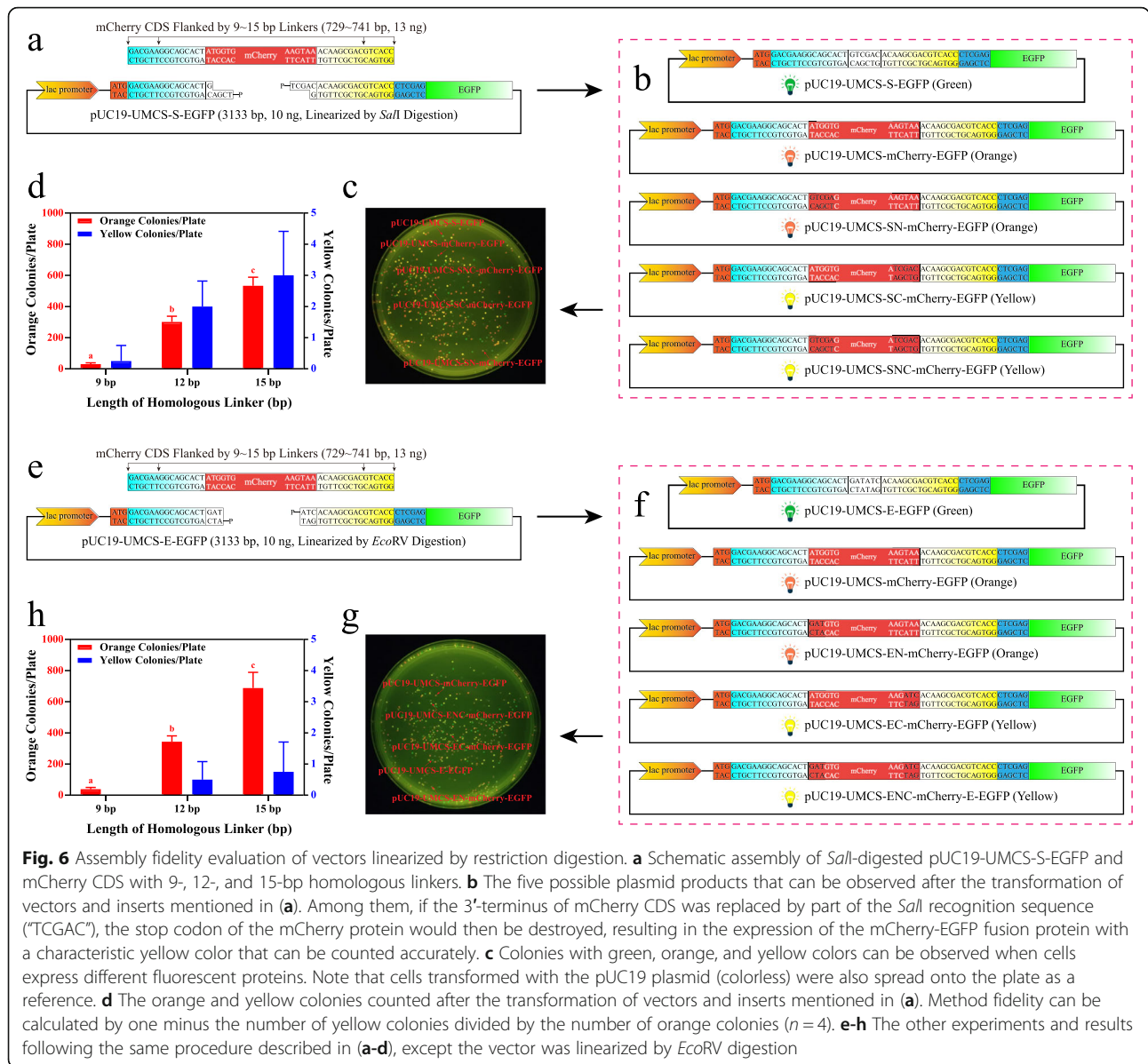


(1) Although high fidelity enzymes such as Q5, Phusion, or KOD have extremely low mismatch rates [27, 28] during the reaction, PCR products could still suffer from random errors introduced by DNA polymerase. Moreover, the possibility of introducing random mutations during the PCR process increases with PCR cycles and the length of vectors, and it is laborious to verify by sequencing. (2) Some vectors are difficult to amplify by PCR. For example, the GC content of the vector is too high or too low, or the vector contains too many repetitive sequences, which will affect PCR and even render the reaction impossible to continue [29–31]. (3) As previously mentioned, if vectors cannot be linearized by PCR or fidelity is vital for the experiment, the methods shown in Fig. 6 must be applied. Under such circumstances, only the sequences flanking the linearization site are used as homologous linkers to ensure the versatility of UMCS. However, when the flanking sequences (15 bp) were used as homologous linkers, the post-digestion residual bases then served as nonhomologous sequences. This means the residual bases might displace part of the insert and cause mutations at its 5'-terminus, 3'-terminus (confirmed during our pilot study), or both. The frequency of such events is a major concern for UMCS applications, since the UMCS is only universal when this frequency is low enough.

Therefore, *SalI* (Fig. 6a-d) and *EcoRV* (Fig. 6e-h) digestion were taken as sticky- and blunt-end examples to evaluate sequence replacement between insert and residual bases after assembly. As shown in Fig. 6, when the mCherry CDS was used as an insert, five possible

products could be expected: (1) if the assembly fails, the colony will appear green (only EGFP expression); (2) if the fragments assembled successfully and the mCherry CDS remains intact, the colony will appear orange (only mCherry expression); (3) if assembled successfully, but the 5'-terminus of the mCherry CDS is partially replaced, the colony will appear orange (only mCherry expression); (4) if assembled successfully, but the 3'-terminus of the mCherry CDS is partially replaced, the colony appears yellow for the loss of the mCherry stop codon (mCherry-EGFP fusion protein is expressed); and (5) if assembled successfully, while both the 5'- and 3'-terminus of the mCherry CDS are partially replaced, the colony also appears yellow (mCherry-EGFP fusion protein is expressed). For subcloning, we assume that the 5'- and 3'-terminus of the insert will be replaced at the same frequency, then the frequency can be derived by calculating the number of yellow colonies against the orange ones. Thus, Fig. 6d strongly suggests that the frequency of mutation at the 3'-terminus of the insert is low when *SalI* is used for vector linearization. In fact, less than 0.75% of positive clones contained sequence replacement at the 3'-terminus or at both the 5'- and 3'-terminus. This indicates that the possibility of having at least one mutated terminus is less than 1.5% when the length of the linker is 12 or 15 bp. Therefore, it will be very unlikely for anyone to pick up a mutated product by chance. Moreover, when *EcoRV* was used for linearization, the frequency of replacement events seemed less than that of *SalI* by having a statistical value of less than 0.3% on both 12- and 15-bp linkers. In





conclusion, after digestion, the residual bases slightly affected the fidelity of assembly (an assumed mechanism of bacterial in vivo assembly with nonhomologous regions is given in Supplementary Fig. S2), but most of the positive colonies were still expected to contain the correct plasmid.

### Construction of larger plasmid based on UMCS

To further demonstrate the versatility of the UMCS, we constructed an additional four vectors containing UMCS: (1) pET24a(+)-UMCS, the UMCS was inserted between the original *Bam*HI and *Xho*I restriction sites, resulting in a plasmid size of 5312 bp; (2) pACT-UMCS and pBind-UMCS, the UMCS was inserted between the original *Bam*HI and *Xba*I restriction sites, the last "C" of

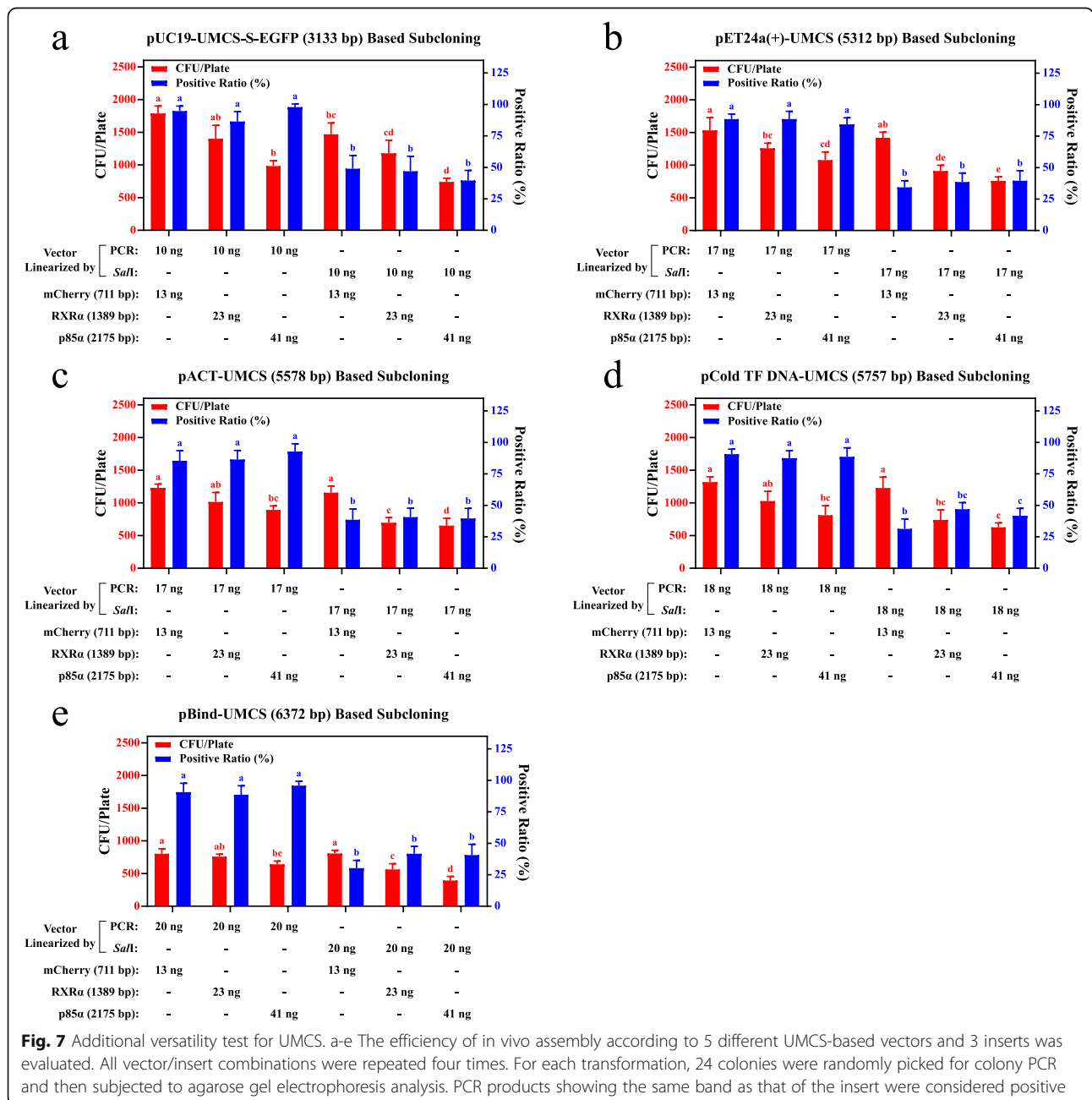
the *Bam*HI recognition sequence was intentionally removed to ensure the correct expression of the insert, resulting in the plasmid size of 5578 and 6372 bp; (3) pCold TF DNA-UMCS, the UMCS was inserted between the original *Nde*I and *Xba*I restriction sites, resulting in a plasmid size of 5757 bp. Together with the pUC19-UMCS-S-EGFP that was constructed previously, we examined the transformation efficiency of cloning the mCherry (711 bp), RXR $\alpha$  (Retinoid X Receptor Alpha, 1389 bp), and p85 $\alpha$  (Phosphoinositide 3-kinase p85 $\alpha$ , 2175 bp) CDS into these vectors. During these experiments, each CDS plus a 15-bp homologous linker region (as shown in Fig. 3a) at both termini was amplified as insert, each transformation was repeated four times, and colony PCR was performed on a panel of 24 colonies

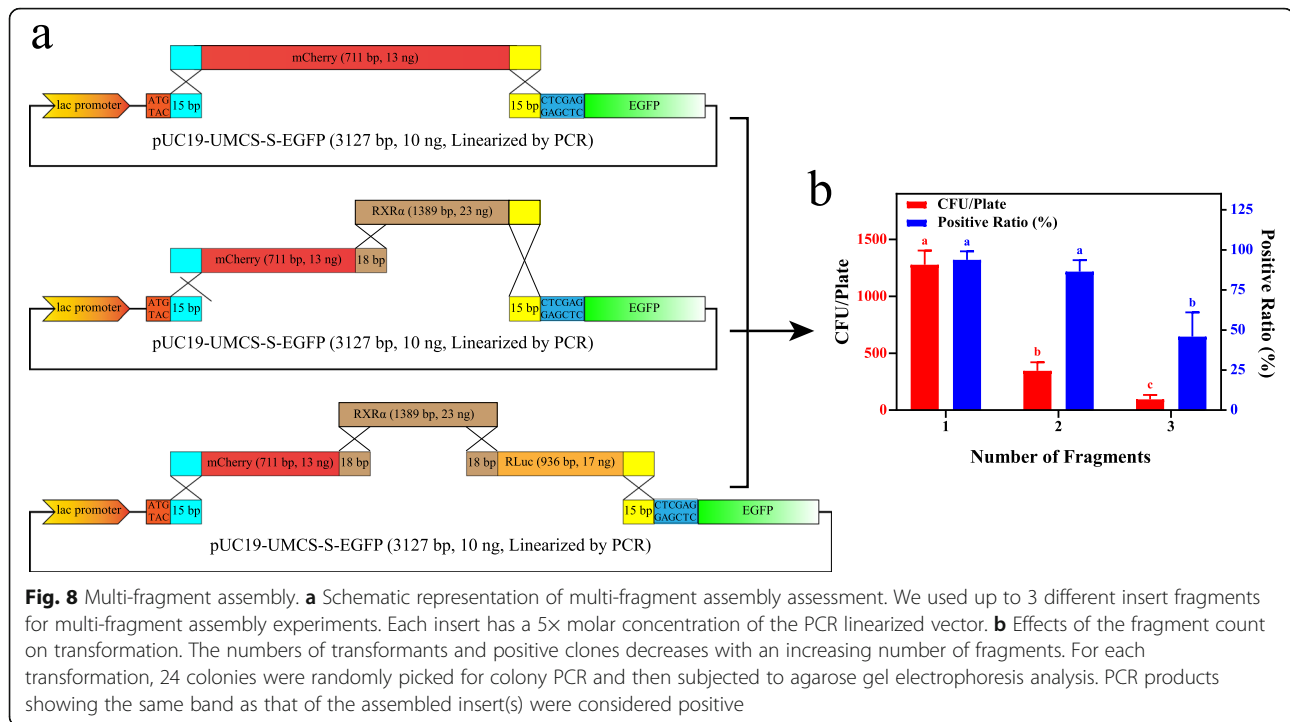
randomly selected from each transformation, followed by agarose gel electrophoresis analysis to identify the positive colonies and the positive ratio accordingly. The corresponding results are shown in Fig. 7. In summary, (1) regardless of the linearization method we tried, the yield of transformants decreased as the size of the final plasmid increased; (2) if PCR was used for vector linearization, no apparent correlations between the positive ratio and the size of the plasmid were observed, and the positive ratios were maintained at a high level (> 85%) in all experiments; and (3) when *SalI* digestion was used, the positive ratio decreased with the increasing size

of the final product. However, in our study, this ratio was always maintained at more than 1/3, which is enough for screening positive clones by PCR.

**Multi-fragment assembly**

Similar to the single-fragment experiments, we used mCherry, RXR $\alpha$ , and RLuc (Renilla Luciferase) for multi-fragment assembly test. As presented in Fig. 8a, the 5'-terminus of the first fragment and the 3'-terminus of the last fragment must have homologous sequences corresponding to the UMCS. The linkers of other fragments needing to be assembled can be





designed according to previous studies [6]. Since multi-fragment assembly significantly reduces the yield of transformants and positive clones (Fig. 8b), PCR as a linearization method or dephosphorylation after enzyme digestion is recommended. Meanwhile, further increasing the amount of DNA and the number of component cells might be necessary for the assembly of more than 3 insert fragments.

## Discussion

In most cases, inserts are required to be placed somewhere within a MCS to function properly. The UMCS-based cloning, like the MCS-based cloning, is not a seamless technique because it relies on synthetic homologous regions to work. However, it provides key advantages over current plasmid cloning methods:

First, we successfully constructed tens of plasmids with single or multiple insertions in a very efficient way by using this method. Although we used DH5α cells throughout this study, other strains such as Top10, BL21(DE3), XL-10 Gold and JM109(DE3) were also tested (as shown in Supplementary Figure S3 and Supplementary Table S2) and had similar success, highlighting the versatility of this protocol.

Second, the *DpnI* enzyme, which is widely used in other reports [6, 7, 20, 26, 32] for the digestion of template DNA, is not necessary when using our protocol because only a few nanograms of linearized vector DNA and a 5× molar concentration of the insert fragment are needed. Thus, a single tube of carefully stored linearized

vector or insert can be used several times without losing efficiency. Therefore, the cost of applying this protocol is well controlled.

Third, the UMCS protocol contains mostly mixing steps when both the vector and insert are ready, making it extremely automation friendly. Instruments for high-throughput transformations are much easier to design and build according to this approach, making this protocol promising in modern molecular biology studies.

Finally, and especially, because of the consistency of homologous linkers and the associated primers, any linearized UMCS-based vector and insert can be mixed directly and is ready for transformation. This is important for those who need to optimize protein expression or switch epitope tags by cloning the same insert into different vectors. In addition, the protocol uses only purified DNA, which means the high-fidelity DNA polymerases used to amplify vectors or inserts are expected to have little or no impact on the transformation. This further guarantees the versatility of UMCS because dedicated vectors or inserts are easier to share between researchers from laboratory to laboratory, regardless of their amplification conditions and without additional PCR steps.

In this study, by comparing different linearization methods, we found that even though the number of transformants was occasionally higher when using the digestion method, the yield of positive clones was more than 20% lower than that of PCR (Fig. 4). Interestingly, the DNA ligase activity inside *E. coli* has shown a great

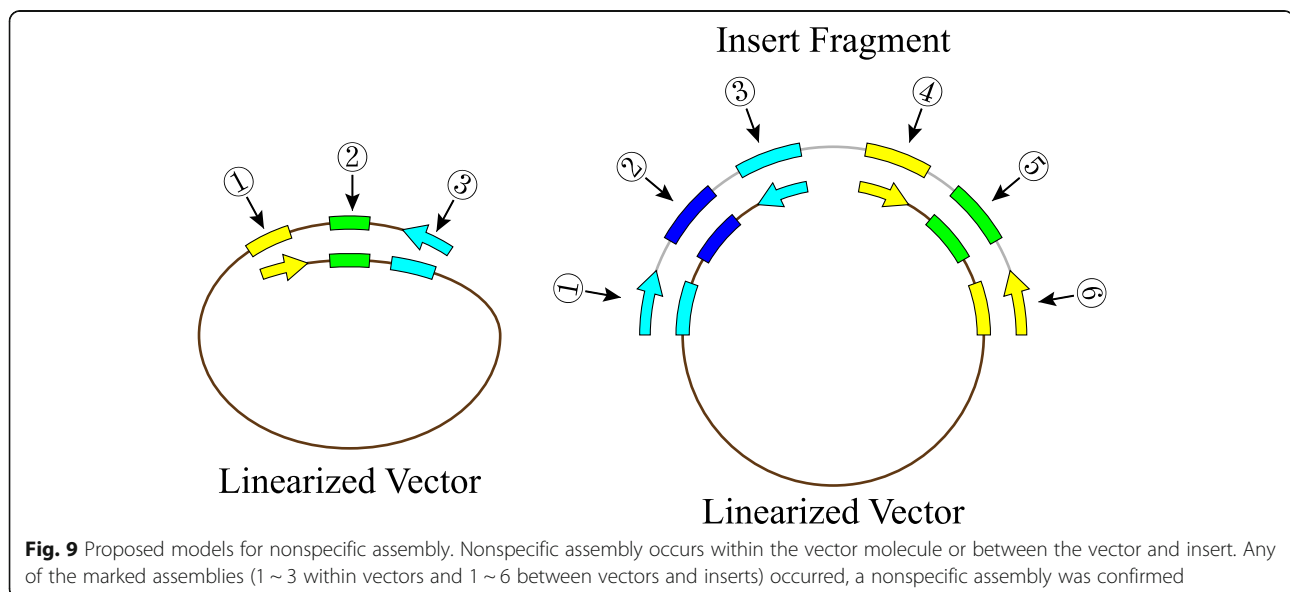


impact on the positive ratio. When comparing Fig. 4a and 4c, we found that when an 18-bp linker was used, the colonies formed after transformation were  $1525 \pm 165$  and  $1483 \pm 121$  without a significant difference, while the positive ratio was reduced significantly from  $96.5 \pm 1.9\%$  to  $75.4 \pm 4.2\%$ , implying the linearized vector with the phosphorylated 5'-terminus is more easily recircularized than the vector without phosphorylation. Additionally, recircularization by DNA ligase is more effective for vectors with sticky ends, which explains why the blunt-ended vector yields a higher positive ratio than the vector with sticky ends, and the PCR-amplified vectors work the best for in vivo assembly. Thus, if the vector itself is small enough or the fidelity of PCR is not a vital concern, PCR is a better choice for linearization, whereas if the vector is large enough that the random mutation introduced by PCR is problematic or the vector is simply resistant to PCR amplification, linearization with the blunt-end enzyme is also an effective strategy. Finally, if no suitable blunt-end enzymes' recognition sequence could be found, digestion by sticky-end enzymes is still a feasible alternative. Furthermore, we believe that if restriction digestion was used for linearization, dephosphorylation might be helpful to further enhance transformation efficiency, especially for multi-fragment assembly [33].

During our research, nonspecific assembly within vectors was observed occasionally. In rare cases, nonspecific assembly between the vector and insert could also be identified. As shown in Fig. 9, we have suggested three nonspecific assembly models within the vector molecule and six models between the vector and insert. When linearized pUC19-UMCS-S-EGFP or pUC19-UMCS-E-

EGFP was involved, the results of nonspecific assembly usually appeared as nonfluorescent (the fluorescent protein or promoter sequence was lost or mutated) colonies. Such cases usually made up less than 5% of the total colonies, regardless of the linearization method. It was also reported previously [34] that the PCR products can be transformed directly and circularized within the cell. Therefore, nonspecific assembly and direct transformation of PCR products might be the two reasons for positive ratios seldom exceeding 97%, even when PCR products purified by agarose gel electrophoresis were used.

Conley et al. [35] proposed a mechanism for bacterial in vivo assembly: double-stranded linear DNA molecules are the substrate for 3'-5'-exonuclease action by ExoIII, which yields single-stranded homologous DNA sequences. After annealing at homologous sites, the vector and insert then circularize to form a plasmid molecule with gaps. The gaps are subsequently filled by DNA polymerase I and ligases as a DNA repair mechanism, thus completing plasmid formation. In this model, non-homologous parts of the ssDNA termini will be displaced by newly-repaired duplex after recombination. However, it is hard to explain solely by this assumption why the intervening nonhomologous sequence (such as post-digestion residual bases) between homologous regions and linear DNA termini still exists as part of the final product in some cases. According to our experiments, such events happened with a probability of around 1.5 and 0.3% when using *SalI* or *EcoRV* digestion, respectively. Although this phenomenon still favors us, how cells decide which sequence to keep within the final plasmid remains unclear. From this perspective,



despite the significant improvements we have made upon understanding the recombination mechanism [36, 37], the picture is still far from complete.

In some cases, the UMCS might affect the function or expression of the insert. For example, if the vector has a start codon located upstream of the UMCS (as shown in Fig. 3 and also typical in pET system for the expression of fusion proteins with N-terminal tags), the sequence of Linker-F (see Fig. 2 for detailed information) will be translated as a short peptide of “DEGST”. If such peptide impairs the insert’s function or expression, it must be removed when constructing the plasmid. The same strategy also applies when the expression of fusion proteins with C-terminal tags (stop codon located downstream of the UMCS) was affected by the sequence of Linker-R (translated as “TSDVT”). Consequently, alternative “half-seamless” or “full-seamless” strategies presented in Fig. 10 can be established. Among them, the “half-seamless” strategies (Fig. 10a and b) for 5’- or 3’-end seamless cloning still simplify the primer design, which will save

time and cost of the experiments and reduce the system complexity.

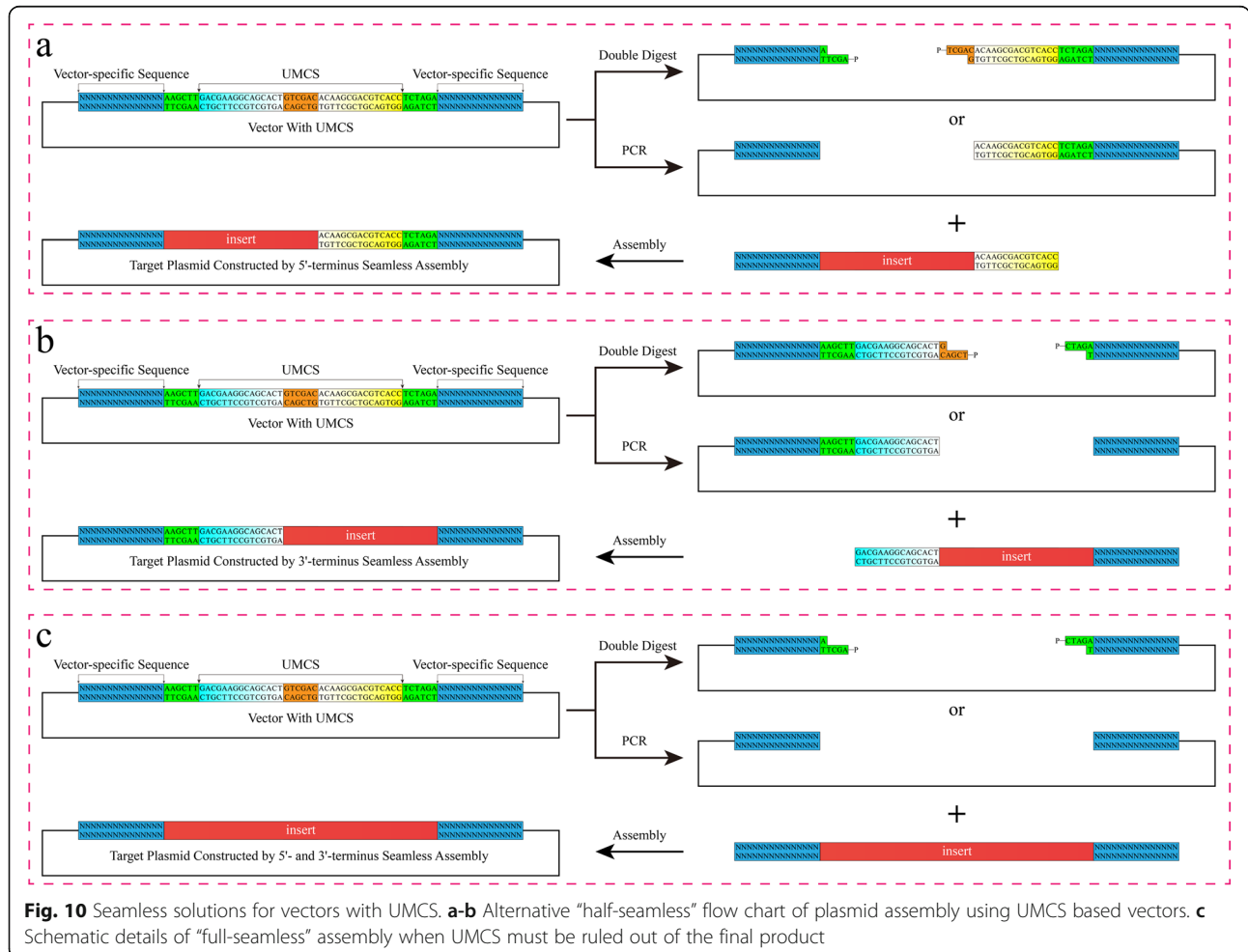
**Conclusions**

In this study, we developed and standardized a digestion-ligation-free approach for rapid and efficient plasmid cloning. The method utilizes a unified MCS design scheme, which simplifies primer design and the following PCR steps. Most importantly, it exploits the in vivo assembly of *E. coli* and significantly reduces hands-on time in the preparation of transformation to just seconds while maintaining excellent reliability.

**Materials and methods**

**Bacterial strains and reagents**

Chemically competent cell strains of *E. coli* DH5α, Top10, BL21(DE3), XL10-Gold, and JM109(DE3) were obtained from Weidi Biotechnology (Shanghai, China) with transformation efficiency of ~ 5 × 10<sup>8</sup>, ~ 5 × 10<sup>8</sup>, ~ 10<sup>7</sup>, ~ 2 × 10<sup>9</sup>, and ~ 10<sup>8</sup> CFU/μg pUC19, respectively. Bacteria containing plasmids were cultured in LB (Luria-



**Fig. 10** Seamless solutions for vectors with UMCS. **a-b** Alternative “half-seamless” flow chart of plasmid assembly using UMCS based vectors. **c** Schematic details of “full-seamless” assembly when UMCS must be ruled out of the final product

Bertani) medium with appropriate antibiotics (ampicillin or kanamycin at 100 or 50 µg/mL, respectively), and IPTG (Isopropyl β-D-1-thiogalactopyranoside, 100 µM) was added when necessary. The SanPrep Column Plasmid Mini-Preps Kit (Sangon Biotech, Shanghai, China) was used for plasmid extraction from bacteria, the SanPrep Column PCR Product Purification Kit (Sangon Biotech) was used for PCR product purification, and the SanPrep Column DNA Gel Extraction Kit (Sangon Biotech) was used for PCR product or enzyme-digested DNA purification after agarose gel electrophoresis, all following the manufacturer's instructions. Both the 4S Red Plus (Sangon Biotech) stained 1% agarose gel and DNA Marker (250 ~ 10,000 bp, Sangon Biotech) were used for all gel electrophoresis experiments. FastDigest *Bam*HI, *Eco*RV, *Nde*I, *Sal*I, *Xba*I, *Xho*I, and T4 DNA ligase from Thermo Fisher Scientific (Shanghai, China) were used for vector construction or linearization. Q5® High-Fidelity DNA Polymerase from NEB (Ipswich, MA, USA) was used to amplify the vector and insert, while Taq PCR Master Mix (Sangon Biotech) was used for colony PCR.

#### Primers and plasmids

Primers were designed using Oligo 7, the  $T_m$  of each primer was calculated according to the "Nearest Neighbor" method. All primers used in this study were designed then synthesized (as shown in Supplementary Table S3) and obtained from Sangon Biotech. The pUC19 plasmid was purchased from Weidi Biotechnology. The pET24a(+) vector was purchased from Invitrogen (Carlsbad, CA, USA). pCold TF DNA was purchased from Takara (Beijing, China). pACT and pBind were purchased from Promega (Madison, WI, USA). pCDNA3.1(+)-EGFP, pCDNA3.1(+)-mCherry, pCMV-RXRα and pCMV-p85α were all gifts from Xiaokun Zhang.

#### PCR conditions and product analysis

Unless otherwise stated, 50 µL PCR was performed using Q5 for vector and insert amplification. The PCR conditions were: 0.02 U/µL polymerase, 0.5 µM each primer, 200 µM dNTPs, and 1 ng prelinearized template DNA. According to the following protocol: 98 °C (30 s) → [98 °C (10 s) → 60 °C (30 s) → 72 °C (2500 bp/min, 20 cycles for vector, 30 cycles for insert)] → 72 °C (2 min) → 10 °C (hold). After amplification, the PCR products were purified and recovered for further use. Additionally, 20 µL PCR was performed using Taq PCR Master Mix for each colony PCR following the manufacturer's instructions. The corresponding PCR condition was: 95 °C (30 s) → [95 °C (15 s) → 55 °C (15 s) → 72 °C (1500 bp/min, 25 cycles)] → 72 °C (2 min) → 10 °C (hold). Agarose gel electrophoresis was used for identifying

positive clones. Colonies that showed the same band as the insert were considered positive.

#### Transformation evaluation

Unless otherwise stated, a 1 µL (4 µL for BL21(DE3)) mixture of linearized vector and insert was transformed into 25 µL (100 µL for BL21(DE3)) of competent cells, and then the mixture was first incubated for 25 min on ice, followed by heat shock at 42 °C for 60 s and transferred to ice for 2 min. After adding 1000 µL of LB medium, cells were subsequently incubated at 37 °C and shaking vigorously (220 rpm) for 45 min. After incubation, cells were centrifuged at 6000 rpm (~ 2500 g) for 1 min, pellets were resuspended in 75 µL of fresh LB medium then spread onto LB agar plates containing antibiotics and IPTG (only for pUC19-based plasmids that express fluorescent protein) for 16 h at 37 °C. Colonies were manually counted, and the positive ratio was assessed by colony PCR (with the "UMCS-Seq-F + UMCS-Seq-R" primer pair) or counting under a 485 nm light source through a 535 nm filter after an additional incubation at 4 °C for 48 h. Part of the positive clones were further confirmed by Sanger sequencing (Sangon Biotech).

#### Data analyses

Data are expressed as the mean ± SD from four or more experiments. Statistical analyses were performed using one-way ANOVA with the comparison methods of the Tukey test (equal variances) or the Games-Howell test (unequal variances). Data sharing a letter in the same group were not considered to be significantly different at the 5% level.

#### Abbreviations

MCS: Multiple Cloning Site; PCR: Polymerase Chain Reaction; IVA: In Vivo Assembly; ELIC: Exonuclease and Ligation-Independent Cloning; LIC: Ligation-Independent Cloning; SLIC: Sequence and Ligation-Independent Cloning; HAC: Homologous Alignment Cloning; SLICE: Seamless Ligation Cloning Extract; IPTG: Isopropyl β-D-1-thiogalactopyranoside; EGFP: Enhanced Green Fluorescent Protein; RXRα: Retinoid X Receptor Alpha; p85α: Phosphoinositide 3-kinase p85α; RLuc: Renilla Luciferase; LB medium: Luria-Bertani medium

#### Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12896-021-00679-6>.

**Additional file 1 Table S1.** An overview of some existing methods developed for plasmid cloning.

**Additional file 2 Table S2.** A selection of UMCS based cloning for XL-10 Gold and JM109(DE3).

**Additional file 3 Table S3.** List of the primers and oligonucleotides used throughout this study.

**Additional file 4 Figure S1.** Effects of insert/vector ratio on assembly efficiency for vector linearized by *Eco*RV digestion.

**Additional file 5 Figure S2.** Assumed mechanism of bacterial in vivo assembly with nonhomologous regions.

**Additional file 6 Figure S3.** Efficiency evaluation of UMCS based subcloning for Top10 and BL21(DE3).

**Additional file 7.** Supplementary image file. Vector format of Figure S1, S2, S5.

**Additional file 8.** Supplementary image file. Uncropped original image of Fig. 6c.

**Additional file 9.** Supplementary image file. Uncropped original image of Fig. 6g.

**Additional file 10.** Supplementary data file. Raw data points as used for Figs. 4, 5, 6, 7, 8, S1, S3.

**Additional file 11.** Supplementary protocol description. Standard protocols for UMCS based cloning.

#### Acknowledgements

The authors thank professor Xiaokun Zhang and his team for providing vital materials.

#### Authors' contributions

FC and YYL conceived and supervised the study. FC, YYL and YLY performed the DNA amplification, purification, and transformation. JD, JLH, carried out the evaluation of transformation efficiency. FC and JL analyzed the data; FC drafted and finalized the manuscript. YYL, YLY reviewed and commented on the manuscript. All authors read and approved the final manuscript.

#### Funding

This work was financially supported by Natural Science Foundation of Fujian Province of China (2018J01464).

#### Availability of data and materials

The datasets used and analyzed in the current study are available from the corresponding author on reasonable request or as in supplementary files.

#### Ethics approval and consent to participate

Not applicable.

#### Consent for publication

Not applicable.

#### Competing interests

The authors declare that they have no competing interests.

Received: 28 November 2020 Accepted: 17 February 2021

Published online: 15 March 2021

#### References

- Tao QZ, Zhang HB. Cloning and stable maintenance of DNA fragments over 300 kb in *Escherichia coli* with conventional plasmid-based vectors. *Nucleic Acids Res.* 1998;26(21):4901–9. Available from: <https://doi.org/10.1093/nar/26.21.4901>.
- Cohen SN. DNA cloning: a personal view after 40 years. *Proc Natl Acad Sci.* 2013;110(39):15521–9. Available from: <https://doi.org/10.1073/pnas.1313397110>.
- Lesley SA, Wilson IA. Protein production and crystallization at the joint center for structural genomics. *J Struct Funct Genomics.* 2005;6:71–9. Available from: <https://doi.org/10.1007/s10969-005-2897-2>.
- Chao R, Yuan Y, Zhao H. Recent advances in DNA assembly technologies. *FEMS Yeast Res.* 2015;15(1):1–9. Available from: <https://doi.org/10.1111/1567-1364.12171>.
- Green M, Sambrook DWJ. *Molecular cloning: a laboratory manual.* 4th ed; 2012.
- Javier GN, Jake WF, Ingo HG. IVA cloning : a single-tube universal cloning system exploiting bacterial in vivo assembly. *Sci Rep.* 2016;6(1):1–12. Available from: <https://doi.org/10.1038/srep27459>.
- Li C, Wen A, Shen B, et al. FastCloning: a highly simplified, purification-free, sequence- and ligation-independent PCR cloning method. *BMC Biotechnol.* 2011;11(1):92. Available from: <https://doi.org/10.1186/1472-6750-11-92>.
- Koskela EV, Frey AD. Homologous recombinatorial cloning without the creation of single-stranded ends: exonuclease and ligation-independent cloning (ELIC). *Mol Biotechnol.* 2015;57(3):233–40. Available from: <https://doi.org/10.1007/s12033-014-9817-2>.
- Li MZ, Elledge SJ. Harnessing homologous recombination in vitro to generate recombinant DNA via SLIC. *Nat Methods.* 2007;4(3):251–6. Available from: <https://doi.org/10.1038/nmeth1010>.
- Tan L, Strong EJ, Woods K, et al. Homologous alignment cloning: a rapid, flexible and highly efficient general molecular cloning method. *PeerJ.* 2018; 6:e5146. Available from: <https://doi.org/10.7717/peerj.5146>.
- Jeong JY, Yim HS, Ryu JY, et al. One-step sequence- and ligation-independent cloning as a rapid and versatile cloning method for functional genomics studies. *Appl Environ Microbiol.* 2012;78(15):5440–3. Available from: <https://doi.org/10.1128/AEM.00844-12>.
- Zhu B, Cai G, Hall EO, et al. In-fusion assembly: seamless engineering of multidomain fusion proteins, modular vectors, and mutations. *BioTechniques.* 2007;43(3):354–9. Available from: <https://doi.org/10.2144/000112536>.
- Walhout AJ, Temple GF, Brasch MA, et al. GATEWAY recombinational cloning: application to the cloning of large numbers of open reading frames or ORFeomes. *Methods Enzymol.* 2000;328:575–92. Available from: [https://doi.org/10.1016/S0076-6879\(00\)28419-x](https://doi.org/10.1016/S0076-6879(00)28419-x).
- Motohashi K. Evaluation of the efficiency and utility of recombinant enzyme-free seamless DNA cloning methods. *Biochem Biophys Rep.* 2017;9: 310–5. Available from: <https://doi.org/10.1016/j.bbrep.2017.01.010>.
- Motohashi K. A simple and efficient seamless DNA cloning method using SLiCE from *Escherichia coli* laboratory strains and its application to SLiP site-directed mutagenesis. *BMC Biotechnol.* 2015;15(47) Available from: <https://doi.org/10.1186/s12896-015-0162-8>.
- Gibson DG, Young L, Chuang RY, et al. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Methods.* 2009;6(5):343–5. Available from: <https://www.nature.com/articles/nmeth.1318>.
- Engler C, Gruetzner R, Kandzia R, et al. Golden gate shuffling: a one-pot DNA shuffling method based on type IIs restriction enzymes. *PLoS One.* 2009;4(5):e5553. Available from: <https://doi.org/10.1371/journal.pone.0005553>.
- Jones DH, Howard BH. A rapid method for recombination and site-specific mutagenesis by placing homologous ends on DNA using polymerase chain reaction. *Biotechniques.* 1991;10(1):62–6.
- Bi X, Liu LF. recA-independent and recA-dependent intramolecular plasmid recombination: differential homology requirement and distance effect. *J Mol Biol.* 1994;235(2):414–23. Available from: <https://doi.org/10.1006/jmbi.1994.1002>.
- Benoit RM, Ostermeier C, Geiser M, et al. Seamless insert-plasmid assembly at high efficiency and low cost. *PLoS One.* 2016;11(4):e0153158. Available from: <https://doi.org/10.1371/journal.pone.0153158>.
- Huang F, Spangler JR, Huang AY. In vivo cloning of up to 16 kb plasmids in *E. coli* is as simple as PCR. *PLoS One.* 2017;12(8):e0183974. Available from: <https://doi.org/10.1371/journal.pone.0183974>.
- Zhu D, Zhong X, Tan R, et al. High-throughput cloning of human liver complete open reading frames using homologous recombination in *Escherichia coli*. *Anal Biochem.* 2010;397(2):162–7. Available from: <https://doi.org/10.1016/j.jab.2009.10.018>.
- Zhang S, Zubay G, Goldman E. Low-usage codons in *Escherichia coli*, yeast, fruit fly and primates. *Gene.* 1991;105(1):61–72. Available from: [https://doi.org/10.1016/0378-1119\(91\)90514-C](https://doi.org/10.1016/0378-1119(91)90514-C).
- Chen DQ, Texada DE. Low-usage codons and rare codons of *Escherichia coli*. *Gene Ther Mol Biol.* 2006;10:1–12.
- Zhang Y, Werling U, Edlmann W. SLiCE: a novel bacterial cell extract-based DNA cloning method. *Nucleic Acids Res.* 2012;40:e55. Available from: <https://doi.org/10.1093/nar/gkr1288>.
- Kostylev M, Otwell AE, Richardson RE, et al. Cloning should be simple: *Escherichia coli* DH5 $\alpha$ -mediated assembly of multiple DNA fragments with short end homologies. *PLoS ONE.* 2015;10(9):e0137466. Available from: <https://doi.org/10.1371/journal.pone.0137466>.
- Potapov V, Ong JL. Examining sources of error in PCR by single-molecule sequencing. *PLoS ONE.* 2017;12(1):e0169774. Available from: <https://doi.org/10.1371/journal.pone.0169774>.
- McInerney P, Adams P, Hadi MZ. Error rate comparison during polymerase chain reaction by DNA polymerase. *Mol Biol Int.* 2014;2014:287430. Available from: <https://doi.org/10.1155/2014/287430>.
- Henke W, Herdel K, Jung K, et al. Betaine improves the PCR amplification of GC-rich DNA sequences. *Nucleic Acids Res.* 1997;25(19):3957–8. Available from: <https://doi.org/10.1093/nar/25.19.3957>.

30. Kumar S. Molecular cloning and expression of high GC-rich novel tumor suppressor gene HIC-1. *Mol Biotechnol.* 2014;56(11):1040–8. Available from. <https://doi.org/10.1007/s12033-014-9783-8>.
31. Riet J, Ramos LRV, Lewis RV, et al. Improving the PCR protocol to amplify a repetitive DNA sequence. *Genet Mol Res.* 2017;16(3):gmr16039796. Available from. <https://doi.org/10.4238/gmr16039796>.
32. Li J, Li C, Xiao W, et al. Site-directed mutagenesis by combination of homologous recombination and DpnI digestion of the plasmid template in *Escherichia coli*. *Anal Biochem.* 2008;373(2):389–91. Available from. <https://doi.org/10.1016/j.ab.2007.10.034>.
33. Jacobus AP, Gross J. Optimal cloning of PCR fragments by homologous recombination in *Escherichia coli*. *PLoS One.* 2015;10(3):e0119221. Available from: <https://doi.org/10.1371/journal.pone.0119221>.
34. You C, Zhang XZ, Zhang YHP. Simple cloning via direct transformation of PCR product (DNA multimer) to *Escherichia coli* and *Bacillus subtilis*. *Appl Environ Microbiol.* 2012;78(5):1593–5. Available from. <https://doi.org/10.1128/AEM.07105-11>.
35. Conley EC, Saunders VA, Jackson V, et al. Mechanism of intramolecular recyclicalton and deletion formation following transformation of *Escherichia coli* with linearized plasmid DNA. *Nucleic Acids Res.* 1986;14(22):8919–32. Available from. <https://doi.org/10.1093/nar/14.22.8919>.
36. Jake FW, Javier GN. In vivo DNA assembly using common laboratory bacteria : a re-emerging tool to simplify molecular cloning. *J Biol Chem.* 2019;294:15271–81 Available from: <https://doi.org/10.1074/jbc.REV119.009109>.
37. Nozaki S, Niki H. Exonuclease III (XthA) enforces in vivo DNA cloning of *Escherichia coli* to create cohesive ends. *J Bacteriol.* 2019;201(5):e00660–18. Available from. <https://doi.org/10.1128/JB.00660-18>.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Ready to submit your research? Choose BMC and benefit from:**

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

**At BMC, research is always in progress.**

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

