

Construction of Expression vectors of (fluorescent) proteins with NBRP vectors Yasunori Sasakura 15th December 2017

Most of NBRP vectors for expression of a protein (including fluorescent protein, fusion protein and Ciona protein; excluding TALENs) share backbone design (Figure 1). The vectors are named “pSPxxxx”. pSP means the vector backbone, and xxxx corresponds to the arbitral ORF inserted in the pSP vector.

(Caution: The above names of NBRP vectors are unofficial according to the nomenclature rule of ascidian community (Stolfi et al., 2015). Their official names are shown in our database.)

The unification of the backbone means that you can use the identical method to modify our vectors. Please assume that you create the vector expressing eGFP in the muscle with the cis element of Troponin I (TnI) and NBRP pSPeGFP vector. When you wish to create the vector expressing Kaede in the muscle with the same cis element, you can use the same primer set, or PCR fragment, of TnI and pSPKaede to create TnI>Kaede construct. The unified system is time and cost saving.

1) ORF switching

When you wish to replace the ORF of pSP vector, you can digest a pSPxxxx vector with BamHI or NotI, and EcoRI to remove xxxx. The recommended choice is NotI and EcoRI, because we can use BamHI site for subcloning cis elements (if BamHI site is not present in the new ORF), and we frequently use BamHI for subcloning cis elements. You may select other restriction sites in the MCS for inserting the ORF; however, please keep in mind that selecting other restriction sites will delete BamHI site which is frequently used for inserting cis elements.

New ORF can be PCR amplified. Previously the PCR amplified ORFs are designed to have BamHI/NotI and EcoRI sites at their ends and are inserted into the BamHI/NotI-EcoRI site of pSP vector with DNA ligase. Recently, In-fusion system is useful for the subcloning because the system is accurate and high efficient. Therefore, we recommend to use the kit.

In-fusion requires 15 base homology arms at both ends of PCR fragment. You can design the primers at the website (<https://www.takara->

bio.co.jp/infusion_primer/infusion_primer_form.php). When an ORF will be inserted into the NotI-EcoRI of pSPxxxx, the primer design of the new ORF will be as follows:

```
GATCCCCTTGCGGCCGCTATGxxxxxxxxxxxxxxxxxxxx  
CGCTCAGCTGGAATTxxxxxxxxxxxxxxxxxxxxxxxxxxxx
```

The xs corresponds to the new ORF. Homology arms are underlined. Red ATG is the start codon of the new ORF.

2) *Cis* element switching

Cis elements can be amplified by PCR from *Ciona* genome or the vectors that have the cis element of your interest. We usually linearize the pSP vectors with BamHI or NotI to insert a cis element. In-fusion system is useful in this case, too. The In-fusion-compatible primers for inserting a cis element into the BamHI site of a pSP vector will be as follows:

```
CGACTCTAGAGGATCxxxxxxxxxxxxxxxxxxxxxxxx  
GGCCGCAAGGGGATCxxxxxxxxxxxxxxxxxxxxxxxx
```

If you like to add the start codon of the endogenous gene, the 3' side primer will be as follows:

```
GGCCGCAAGGGGATCCATxxxxxxxxxxxxxxxxxxxxxxxx
```

Red indicates the start codon.

If the pSPxxxx vector obeys our recommendation suggested in the section 1-1), you can re-use the PCR fragment of the cis element to subclone it into the BamHI of pSPxxxx vector by In-fusion system, because the homology arm matches. If the pSPxxxx vector does not obey the rule, you may have to order another 3' side primers for the subcloning.

When isolating a new cis element whose activity has not been validated in literatures or experiments, we usually start amplifying approximately 2,000 bp fragment upstream from the start codon. Because start codon usually presents in the

1st exon, the fragment includes the region encoding 5'UTR. When the distance between the next gene is shorter than 2,000 bp, please reduce the length of amplicon.

Tips: Some cis elements require longer upstream regions. If a new cis element does not work, please isolate longer fragment. We have experienced that some enhancers are included in the first intron. In this case, addition of the first intron would improve the situation. The requirement of the 1st intron could be predicted by the homology between *C. intestinalis* and *C. savignyi* (that can be found in the JGI Ciona genome browser). The way to introduce enhancers in expression vectors is discussed in the next section.

3) Introduction of enhancers

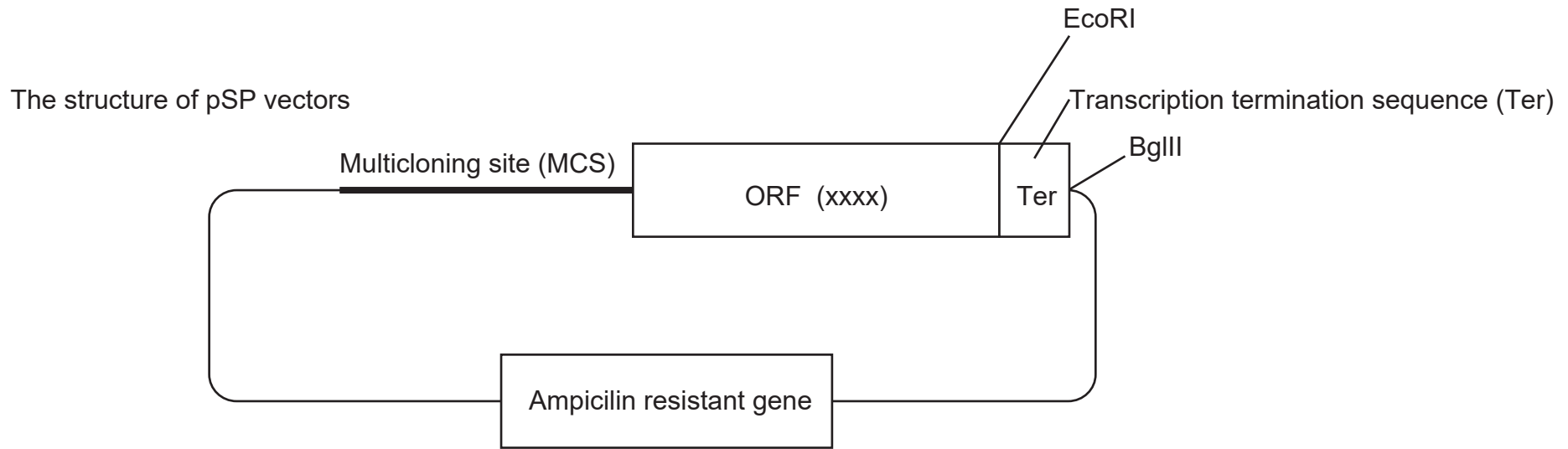
Sometimes we wish to introduce another DNA fragment in your pSP-cis element>xxxx vector. There are two choices of the position of another DNA, the upstream of the cis element or the downstream of ORFs.

If upstream region from the cis element is the good position, you can use restriction sites such as XhoI, Sall or PstI for this purpose (Figure 1). Linearize vectors with one of these restriction site and insert the enhancer by In-fusion reaction. This method is good when enhancers that are insensitive to the position relative to the gene, or you wish to insert a cis element that is too long to isolate by a single PCR reaction.

When the 1st intron is necessary, we usually insert the DNA encoding the intron at the downstream region of the ORF. For this purpose, BglII site that is at the 3' end of the transcription termination sequence (Ter) can be used.

In both cases, the restriction enzymes could not sometimes be used because of the presence of another restriction site in the inserted cis element or ORF. In such cases, you can linearize the vector by inverse PCR.

Figure 1



MCS of pSPnEGFP:

CTCGAGCAGCTGAAGCTTGCATGCCTGCAGGTCGACTCTAGAGGATCCCCTTGCGGCCGCAAGGGTACGGCGCTCAGAAAAAatg

XhoI HindIII SphI PstI SalI XbaI BamHI NotI Start codon of NLS-eGFP

The MCS sequence is not always identical between vectors because some nucleotides are lost during cloning; however, the reading frame is always the same with respect to the restriction sites. For example, all vectors have 3n+1 nucleotides between BamHI site and start codon.