

# **Construction of expression vectors of (fluorescent) proteins with NBRP vectors**

**Yasunori Sasakura 15th December 2017**

**Modified on 9th March 2021**

**Modified on 26th January 2023**

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

(Caution: The above names of NBRP vectors are unofficial according to the nomenclature rule of the 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 another 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 a TnI>Kaede construct. The unified system saves time and money.

## **1) ORF switching**

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

The new ORF can be PCR-amplified. Traditionally 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 the pSP vector with a DNA ligase. Recently, the In-

fusion system has been useful for the subcloning because the system is accurate and highly efficient. Therefore, we recommend using the kit.

In-fusion requires 15 base homology arms at each end of the PCR fragment. You can design the primers at the website ([https://www.takara-bio.co.jp/infusion\\_primer/infusion\\_primer\\_form.php](https://www.takara-bio.co.jp/infusion_primer/infusion_primer_form.php)). When an ORF is to be inserted into the NotI-EcoRI of pSPxxxx, the primer design of the new ORF will be as follows:

```
GATCCCCTTGCGCCGCTATGxxxxxxxxxxxxxxxxxx  
CGCTCAGCTGGAATTxxxxxxxxxxxxxxxxxxxxxxxx
```

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 the *Ciona* genome or from vectors that have the cis elements of your interest. We usually linearize pSP vectors with BamHI or NotI to insert a cis element. The 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 are as follows:

```
CGACTCTAGAGGATCxxxxxxxxxxxxxxxxxxxxxxxx  
GGCCGCAAGGGGATCxxxxxxxxxxxxxxxxxxxxxxxx
```

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

```
GGCCGCAAGGGGATCCATxxxxxxxxxxxxxxxxxxxxxxxx
```

Red indicates the start codon.

If the pSPxxxx vector obeys our recommendation in section 1-1), you can re-use the PCR fragment of the cis element to subclone it into the BamHI of the 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 primer for the

subcloning.

When isolating a new cis element whose activity has not been validated in the literature or in experiments, we usually start by amplifying an approximately 2,000 bp fragment upstream from the start codon. Because the start codon usually presents in the first exon, the fragment includes the region encoding the 5'UTR. When the distance to the next gene is shorter than 2,000 bp, the length of the amplicon you use should be reduced.

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

### **3) Introduction of enhancers**

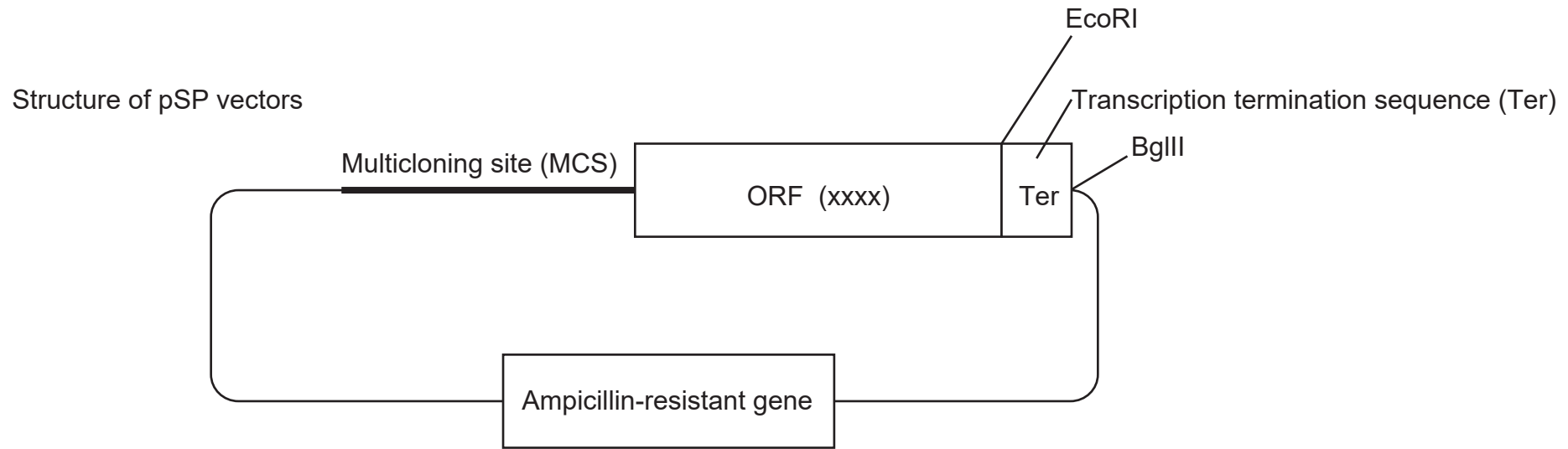
Sometimes we wish to introduce another DNA fragment in your pSP-cis element>xxxx vector. There are two choices for the position of another DNA: upstream of the cis element or downstream of the ORFs.

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

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

In both cases, sometimes the restriction enzymes cannot 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 the BamHI site and the start codon.