# **S1 Appendix**

## **Vector modification, promoter candidate fragment preparation and assay vector preparation**

pMCS-Cypridina Luc vector (Cat#16149, Thermo Fisher SCIENTIFIC) has eight cloning sites and all of them can be cut by six base recognition restriction enzyme. But in that case, some desired promoter candidate region may have that recognition site. It is cut in the middle of DNA fragment. It is inconvenient for ligating. To perform ligation efficiently, we decided to modify the vector to have a Homing endonuclease site that recognizes rare sequences.

Double strand DNA cassette was constructed by annealing the DNA oligo that has SpeI site/I-CeuI site and the DNA oligo that has BamHII site/I-SceI site (DNA oligos were synthesized by Eurofin Genomics). And next, its DNA cassette was digested with SpeI (Cat#R0133S, New England Biolabs)/BamHI-HF (Cat#R3136S, New England Biolabs). On the other hand, pMCS-Cypridina Luc vector (Cat#16149, Thermo Fisher SCIENTIFIC) was digested with SpeI/BamHI-HF and both 5’-ends of vector were dephosphorylated with FastAP Thermosensitive Alkaline Phosphatase (Cat#EF0654, Thermo Fisher SCIENTIFIC).

DNA cassette after digestion with SpeI/BamHI and the vector after digestion with SpeI/BamHI were ligated using Rapid DNA ligation Kit (Cat# K1422, Thermo Fisher SCIENTIFIC). Next, One Shot TOP10 Chemically Competent E. coli (Cat# C404006, Thermo Fisher SCIENTIFIC) was transformed by the ligated vector. These transformed E. coli cells were cultured in large scale and modified vector was extracted using QIAGEN Plasmid Mini Kit (Cat# 12123, QIAGEN). It is named “pMCS-Cypridina Luc wCST”. pMSC-Cypridina Luc wCST was digested with homing endonuclease I-CeuI (Cat#R0699S, New England Biolabs) and I-SceI (Cat#R0694S, New England Biolabs), and both 5’-ends of vector were dephosphorylated with FastAP Thermosensitive Alkaline Phosphatase.

Specific DNA primers were designed for amplification of promoter candidate (DNA primers were synthesized by Eurofin Genomics). Using these primers and Human Genomic DNA (Cat# G1471, Promega) as template, PCR was performed with KOD-Plus-Neo (Cat#KOD-40, TOYOBO). Reaction conditions were followings. Basically 2 step PCR (94℃ 2min->[98℃ 10sec->68℃ 30sec] x 35cycles->68℃ 10min) was performed. PCR of the sample that appeared multi-bands were performed with step-down PCR protocol (94℃ 2min->[98℃ 10sec->68℃ 30sec] x 35cycles->68℃ 10min). As for PCR for the samples in the result of no amplification, 3step PCR protocol (94℃ 2min->[98℃ 10sec->74℃ 30sec] x 5cycles->[98℃ 10sec->72℃ 30sec] x 5cycles->[98℃ 10sec->70℃ 30sec] x 5cycles->[98℃ 10sec->68℃ 30sec] x 25cycles-> 68℃ 10min) was used. After QC with electrophoresis, PCR products (insert DNAs) were digested by I-CeuI/I-SceI. We selected and prepared human GAPDH promoter region as positive control and random sequence (backbone of pMCS-Cypridina Luc vector (201 bases length)) as negative control.

The above prepared vector and insert DNA were ligated using Rapid DNA ligation Kit, and One Shot TOP10 Chemically Competent E. coli was transformed by the ligated vector. These transformed E. coli cells were cultured in large scale and assay vectors were extracted using QIAGEN Plasmid Mini Kit. For quality checking, PCR was performed with assay vector and primers for checking. After this, electrophoresis was performed to confirm the presence of the desired insert DNA fragment in the assay vector. In addition to the self-made control vector, pCMV-Cypridina Luc (Cat # 16150, Thermo Fisher SCIENTIFIC) purchased as a positive control and pMCS-Cypridina Luc w CST with no insert DNA were prepared as a negative control.

## **Assay vector transfection and luciferase assay**

Until the day before doing assay, HEK293T cells were pre-cultured (DMEM (Cat#11965-092, Gibco) + 10% FBS (Cat# 26140-079, Gibco)). At the day before doing assay, pre-cultured HEK293T cells were spread on 96-well plate (Cat#167008, Thermo Fisher SCIENTIFIC), at the number of 1x104 /well in DMEM+10%FBS medium and was incubated at 37℃ CO2 5% for 16-24h. After incubation, assay vector was transfected into HEK293T cells using Turbofect Transfection Reagent (Cat#R0533, Thermo Fisher SCIENTIFIC) according to the kit protocol. To correct the difference of luminescence intensity caused by the difference of transfection efficiency and cell number, pCMV-Red Firefly Luc vector (Cat#16156, Thermo Fisher SCIENTIFIC) was co-transfected together. Vector-transfected HEK293T cells were incubated 37℃ CO2 5% for 24h. After incubation, Cells were lysed using Cell lysis buffer included in Pierce Cypridina-Firefly Luciferase Dual Assay Kit (Cat# 16183, Thermo Fisher SCIENTIFIC). According to the kit protocol, prepared reagent mixture containing D-Luciferin were added to the cell lysate and measurement of luminescence from luciferin-luciferase reaction was performed by luminometer (2030 ARVO X, Perkin Elmer).

**Original vector for assay (before modification)**

**Assay vector (after modification)**

wCST

**I-CeuI**

**I-SceI**

**Vector for correction (co-transfection with assay vector)**



**Outline of experiment (vector modification - construction of assay vector)**

**Assay vector transfection to HEK293T**

correction vector

(pCMV-Red Firefly Luc)

Assay vector

Expression of luciferase protein

HEK293T cell

HEK293T cell

Transfection

Incubation for 37℃ CO2 5% 24h

**Luciferase assay**

(Referred from Thermo Fisher SCIENTIFIC)

