

## **Expression Arrest™ Human retroviral shRNA<sup>mir</sup> individual constructs**

Catalog # RHS1764

The laboratory of Dr. Greg Hannon at Cold Spring Harbor Laboratory (CSHL) has created an RNAi Library comprised of multiple short-hairpin RNAs (shRNAs) specifically targeting annotated human genes. The shRNA Library permits rapid, cost-efficient, loss-of-function genetic screens and rapid tests for genetic interactions to be performed in mammalian cells.

Each shRNA construct has been sequence verified and cloned into a retroviral vector to ensure a match to the target gene. The shRNA expression cassette is carried in a validated Murine Stem Cell Virus (MSCV) backbone. The pSHAG-MAGIC vector (See Figures 1 & 2) can be used both for transient and stable delivery by transfection and for stable delivery using the replication-deficient retrovirus as a delivery method.

### ***shRNA construct storage***

Each Expression Arrest™ Human shRNA construct is shipped as a bacterial culture of *E. coli* (DH10bpir116) in LBR broth with 8% glycerol and chloramphenicol (50ug/ml). Open Biosystems checks all cultures for growth prior to shipment.

Individual constructs are shipped on wet ice. All Expression Arrest™ Human shRNA constructs should be stored at -80°C.

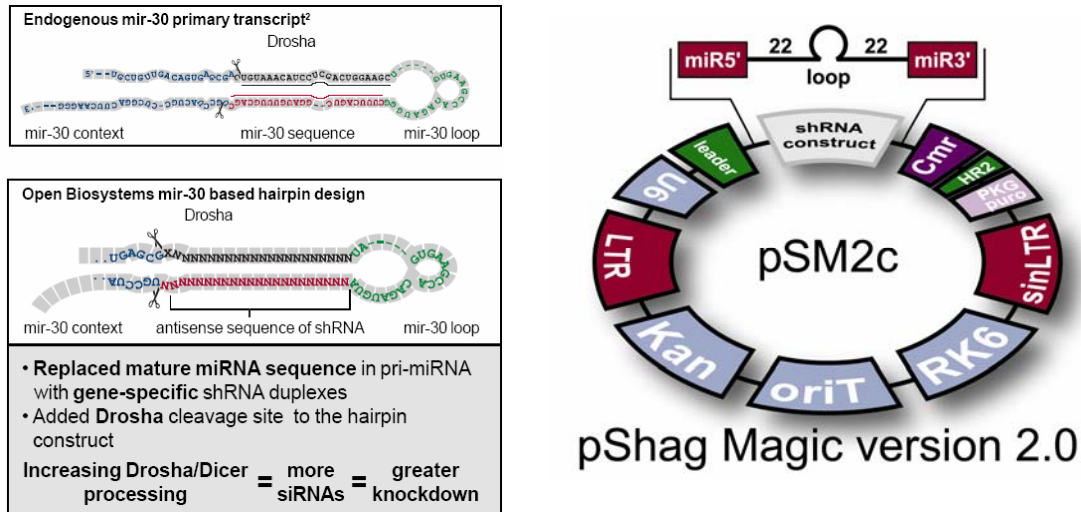
### **Expression Arrest™ Human shRNA collection**

Expression Arrest™ human shRNA library from Open Biosystems is a whole genome RNAi resource and the only choice for transient, stable and *in vivo* RNAi studies. This collection was developed in collaboration with Drs Greg Hannon (CSHL) and Steve Elledge (Harvard). The collection has several unique features that make it a very versatile and efficient tool for RNAi studies including large-scale screens (Paddison *et al* 2004).

#### **These include:**

##### **(1) Unique MicroRNA-30 based hairpin design**

Expression Arrest™ short hairpin RNA constructs are expressed as human microRNA-30 (miR30) primary transcripts. This design adds a Drosha processing site to the hairpin construct and has been shown to greatly increase knockdown efficiency (Boden *et al* 2004). The hairpin stem consists of 22-nt of dsRNA and a 19-nt loop from human miR30. Adding the miR30 loop and 125nt of miR30 flanking sequence on either side of the hairpin results in greater than 10-fold increase in Drosha and Dicer processing of the expressed hairpins when compared with conventional shRNA designs without microRNA. Increased Drosha and Dicer processing translates into greater siRNA/miRNA production and greater potency for expressed hairpins.



**Figure 1:** Expression Arrest™ shRNA are expressed as mir30 primary transcripts

Use of the miR30 design also allowed the use of **'rules-based' designs** for target sequence selection. One such rule is the destabilizing of the 5' end of the antisense strand which results in strand specific incorporation of miRNAs into RISC. The proprietary design algorithm targets sequences in coding regions and the 3'UTR with the additional requirement that they contain greater than 3 mismatches to any other sequence in the human or mouse genomes. Due to the placement of the RNA Polymerase III transcription terminator (four or more thymidines) downstream of the hairpin, each transcript is designed to precisely terminate. RNA Polymerase III terminates on the second thymidine, two uridines remain to create a 2 base overhang. Each shRNA construct has been sequence verified before being cloned into the retroviral vector to ensure a match to the target gene. To assure you the highest possibility of modulating the gene expression level, each gene is represented by multiple shRNA constructs, each covering a unique region of the target gene.

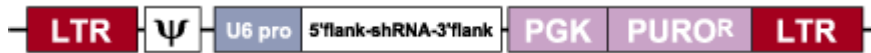
**(2) Versatile vector design**

Expression Arrest™ shRNA are already cloned into the pSHAG-MAGIC2 (pSM2) retroviral vector. This vector has a Murine Stem Cell Virus (MSCV) backbone. Features of the vector that make it a versatile tool for RNAi studies include:

- Ability to perform transfections (transient and stable) or transductions using the replication incompetent retrovirus
- Amenable to *in vitro* and *in vivo* applications
- Puromycin drug resistance marker for selecting stable cell lines
- Molecular barcodes enable complex screening in pools

The pSHAG-MAGIC2 cloning vector is roughly equivalent to pSHAG-MAGIC1 (see Paddison *et al.*, 2004) with a few notable exceptions. First, the cloning strategy has been changed. Previously “PCR-SHAG” was used to clone hairpins by adding the entire hairpin onto the end of a PCR primer. Now a single oligo is used, containing the hairpin

and common 5' and 3' ends, as a PCR template. That is, the oligo is PCR amplified using universal primers that contain *XhoI* (5' primer) and *EcoRI* (3' primer). These PCR fragments are then cloned into the hairpin cloning site of pSHAG-MAGIC2. The mir30-styled hairpins are still expressed from the human U6 promoter. The configuration of pSHAG-MAGIC2 is shown below. The 5' and 3' flanks are derived from 125 bases surrounding the Human miR30 microRNA.



Vector Element	Utility
U6 promoter	RNA generated with 4 uridine overhangs at each 3' end
Retroviral Signaling Sequence	Combined with packaging extract for mammalian cell infection
PGK-Puro	Selection for transfection stability in mammalian cells
Chloramphenicol/Kanamycin	Bacterial selection marker
Homologous recombination sites	Transfer shRNA cassette into new vectors through the MAGIC homologous recombination system
<b>RK6</b>	Conditional origin of replication. Requires the expression of pir1 gene within the bacterial host to propagate

**Table 1: Features of the pSHAG-MAGIC2 Vector**

### Antibiotic Resistance

**pSHAG-MAGIC2 contains 4 antibiotic resistance markers. (See Table 2)**

Antibiotic	Concentration	Utility
Chloramphenicol	50ug/ml	Bacterial selection marker (shRNA insert)
Kanamycin	optional	Bacterial selection marker (vector)
Puromycin		Mammalian selectable marker

**Table 2: Antibiotic Resistances Conveyed by pSHAG-MAGIC2**

### **Culturing protocols and maintenance of pSM2**

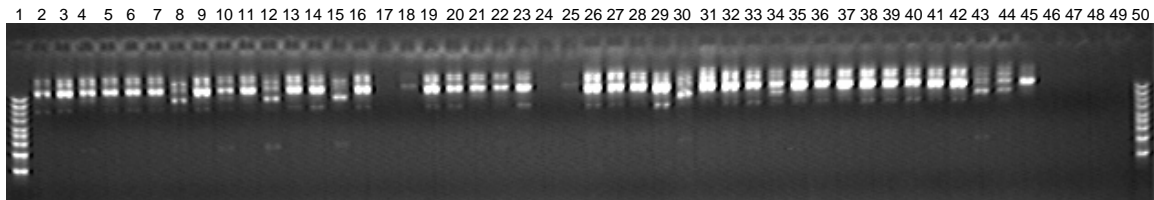
It is well known that viral vectors have a tendency to recombine producing background recombinants. Recombination occurs at the long terminal repeat regions (LTR's). The LTR recombination, which results in loss of most of the plasmid, can confer a growth advantage on the cells. It is therefore critical to maintain careful growth conditions when culturing viral vectors in *E.coli* in order to reduce the number and abundance of background recombinants.

The Expression Arrest™ shRNA Library version 1.3 is constructed in the multi-functional pSM2 vector. This vector allows for both transient and stable gene knockdown via the mechanism of RNA interference. The vector is capable of producing self-inactivating murine-stem-cell-virus (MSCV) particles when used in conjunction with retroviral packaging lines.

pSM2 is a viral vector that produces very little recombinant background product under careful growth and handling conditions. We have observed that greater than 24-hour incubation times increases recombination only slightly.

In order to obtain a good yield of cells in a short period of incubation, rich media (containing 8% glycerol) should be used to culture pSM2 constructs. An incubation period of 14-20 hours at 37°C with aeration is sufficient. It is recommended that the cultures remain frozen at -80°C when not in use. Freeze/thaw cycles do not seem to have any detrimental effect providing the cultures are not incubated at room temperature or higher, for long periods of time.

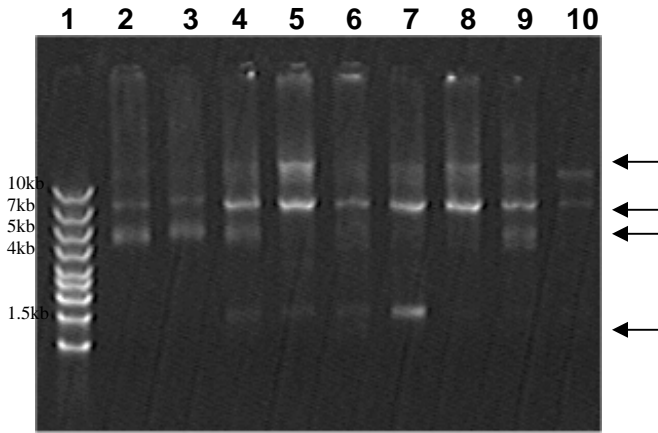
Gel images of plasmid isolated from cultures grown under the above conditions are shown below.



**Figure 2.** 1.5 ml cultures of 42 different shRNA constructs after 20 hours of incubation at 37°C with shaking (~170 rpm). 2X LB media (low-salt) with 8% glycerol was used for culturing. This vector is a stable retroviral vector and shows minimal recombinants. The pSM2 band usually runs around 7kb although it is not uncommon to see a band around 10kb or even around 5 kb. The presence of a faint recombinant band is seen around 1.8 kb in lanes 10 and 12. If the recombinant product is not a significant proportion (over 50%) of your plasmid prep the DNA is still acceptable for transfection since the LTR-LTR recombinant product does not contain the puromycin resistance gene or the shRNA construct.

### Background recombination levels associated with pSM2

Although careful growth conditions were maintained when culturing this set, a small percentage of the whole set (~5%) still shows a low level of recombination. The following gel image is an example of what to expect after plasmid DNA preparation.



**Figure 3:** Lane 1– 10kb marker (10kb, 7kb, 5kb, 4kb, 3kb, 2.5kb, 2kb, 1.5kb, 1kb), Lanes 2 to 10– 10ul of plasmid prep product of nine different shRNA constructs.

A 1 $\mu$ l inoculum of 9 different shRNA constructs were cultured in 1ml of 2XLB medium (low-salt) in a bioblock with aeration by shaking at 200 RPM at 37° C for 16 hours. Plasmids were isolated and run uncut on a 0.9% agarose-TAE gel. The first three arrows from the top point to various forms of the correct plasmid pSM2, which when digested with restriction enzymes produces the correct band size. The last arrow from the top points to the recombinant product (~ 1.8kb). Samples on lanes 4, 5, 6 and 7 show varying levels of recombination. Samples 2, 3, 8, 9 and 10 show minimal to no recombination.

## **Protocols for culturing pSM2 shRNA constructs**

### **2X-LB broth (low-salt) media preparation for plasmid DNA**

Peptone	20 g/L
Yeast Extract	10 g/L
NaCl	5g/L
Chloramphenicol	50ug/ml
*Glycerol	8% for long term storage

*Note: (1) LB media can be used instead of 2XLB*

*(2) \*Glycerol can be omitted from the media if you are culturing for plasmid preparation. If making copies of the constructs for long term storage at  $-80^{\circ}\text{C}$ , 8% glycerol is required.*

### **Culture conditions for individual plasmid preparations**

Most plasmid mini-prep kits recommend a culture volume of 1–10 ml for good yield. For shRNA constructs, 5ml of culture can be used for one mini-prep generally producing from 5–20 ug of plasmid DNA.

1. Upon receiving your glycerol stock(s) containing the shRNA of interest store at  $-80^{\circ}\text{C}$  until ready to begin.
2. To prepare plasmid DNA first thaw your glycerol stock culture and pulse vortex to resuspend any *E. coli* that may have settled to the bottom of the tube.
3. Using a sterile loop or a pipette tip, streak the shRNA culture onto a LB agar plate containing 50 ug /ml Chloramphenicol. Incubate the plate overnight at  $37^{\circ}\text{C}$ . Return the glycerol stock(s) to  $-80^{\circ}\text{C}$ .
4. The following day, pick 1 to 3 colonies from the agar plate and inoculate 6 ml of the 2XLB Chlor<sub>50</sub>. Incubate at  $37^{\circ}\text{C}$  for 16-20 hrs with vigorous shaking (300 rpm).
5. The following day remove 1 ml of the culture and place in a sterile 2-ml sterile microcentrifuge tube. Place this tube at  $4^{\circ}\text{C}$  until the plasmid DNA from the remaining culture has been analyzed. Pellet the remaining 5-ml culture and begin preparation of plasmid DNA. We recommend preparing Ultra-pure DNA to ensure both high-purity and low endotoxin levels (Qiagen Catalog #12123) as required for transfection into eukaryotic cells.

*If you wish to continue at a later time cell pellets can be kept frozen at  $-20^{\circ}\text{C}$  overnight.*

6. Run 3-5ul of the plasmid DNA on a 1% agarose gel. The uncut pSM2 shRNA constructs run at about 7-10kb while the most common product of a recombination event will run at  $\sim 1.5$ -1.8kb. If recombination is present at a significant amount then return to the plate and pick another colony and repeat plasmid preparation. A small amount of recombination is acceptable during transfection since the LTR-LTR recombinant product does not contain the puromycin resistance gene or the shRNA.
7. Prepare an 8% glycerol stock culture using the 1ml of culture you removed prior to plasmid preparation. This culture can be used for future plasmid preparations but it is still recommended you streak isolate and work from a fresh colony. Store at  $-80^{\circ}\text{C}$ .

*Note: Due to the tendency of all viral vectors to recombine we recommend keeping the incubation times as short as possible and avoid subculturing. Return to your original glycerol stock or the colony glycerol stock for each plasmid preparation.*

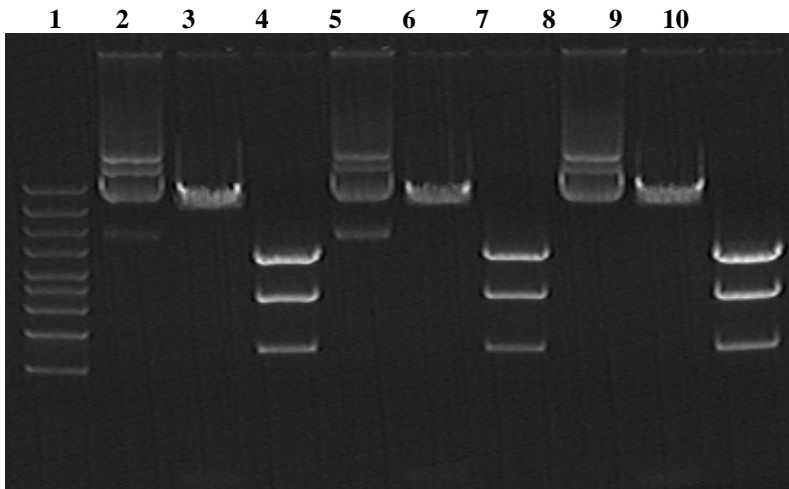
### Restriction Digests of pSM2

You may wish to restriction digest a sample of your plasmid DNA following plasmid DNA preparation. The following is a protocol for dual restriction enzyme digestion using EcoRI and XhoI for quality control of pSM2 vectors (shRNA library and controls). The protocol for HindIII/XbaI digests is exactly the same except replace the EcoRI Buffer with the 10X Buffer 2 and exchange the enzymes used.

1. Using filtered pipette tips and sterile conditions add the following components, in the order stated, to a sterile PCR thin-wall tube.

Sterile, nuclease-free water	14.8µl
Restriction enzyme EcoR1 10X buffer	2µl
BSA (10X, 10mg/ml)	0.2µl
DNA sample 1µg, in water or TE buffer	1µl
Restriction enzyme EcoRI, 20U	1µl
Restriction enzyme XhoI, 20U	1µl
<b>Final volume</b>	<b>20µl</b>

2. Mix gently by pipetting.
3. Incubate in a thermalcycler at 37°C for 2.5 hours to digest then at 70°C for 20 minutes to kill the enzyme.
4. Add 4µl of 6X Loading Dye (or another appropriate DNA loading buffer), and proceed to gel analysis.
5. Load the gel with 20µl of each of the digested samples (a EcoRI/XhoI and HindIII/XbaI) on a 1% agarose gel. Also run 1µl (1µg) of the uncut sample combined with 16µl of water and 3µl of 6x dye alongside the digested samples.
6. The EcoRI/XhoI digest will release the 97-bp insert and leave an approximately 7-kb band. The XbaI/HindIII digests should have 3 bands: 3690bp, 2260bp and 1253bp.



**Figure 4.** The 1% agarose gel above contains -10kb ladder followed by undigested

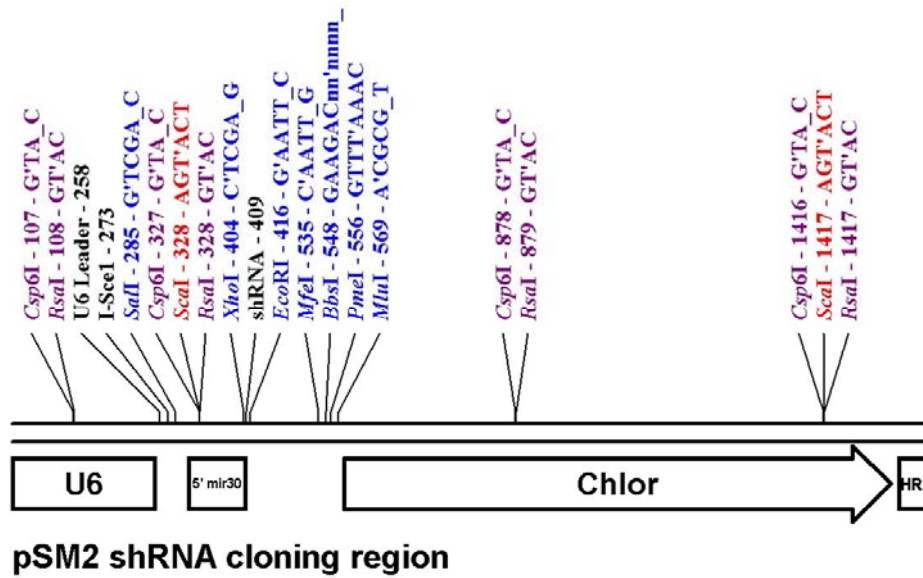


CTTCCCAACCTTACCAGAGGGCGCCCCAGCTGTCCGAAATATTATAAATTATCGCACACATAAAA  
ACCATGCTGTTGGTGTGTCTATTAAATCGGCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTT  
CGCTATTACGCCAGCTGGCGAAAGGGGATGTGCTGCAAGGCGATTAAGTTGGGTAAACGCCAGGG  
TTTTCCCAGTCACGACGTTGTA AAAACGACGGCGCAAGGAATGGTGATGCAAGGAGATGGCGCCC  
AACAGTCCCCCGGCCACGGGGCCTGCCACCATACCCACGCCGAAA CAAGCGCTCATGAGCCCGAA  
GTGGCGAGCCCGATCTTCCCACATCGGTGATGTGCGCGATATAGGCGCCAGCAACCGCACCTGTGG  
CGCCGTGATGCCGGCCACGATGCGTCCGGCGTAGAGGCGATTAGTCCAATTTGTTAAAGACAGG  
ATATCAGTGGTCCAGGCTCTAGTTTGGACTCAACAATATCACCAGCTGAAGCCTATAGAGTACGA  
GCCATAGATAAAAATAAAAGATTTTATTTAGTCTCCAGAAAAGGGGGGAATGAAAGACCCACCT  
GTAGGTTTGGCAAGCTAGCTTAAAGTAACGCCATTTTGCAAGGCATGGAAAATACATAACTGAGAA  
TAGAGAAGTTAGATCAAGGTTAGGAACAGAGAGACAGCAGAATATGGGCCAAAACAGGATATCTG  
TGGTAAGCAGTTCTGCCCCGGCTCAGGGCCAAGAACAGATGGTCCCAGATGCGTCCCAGCCCT  
CAGCAGTTTCTAGAGAACCATCAGATGTTTTCCAGGGTGCCCCAAGGACCTGAAATGACCCTGTGC  
CTTATTTGAACTAACCAATCAGTTCGTTCTCGCTTCTGTTCTGCGCGCTTCTGCTCCCCGAGCTC  
AATAAAAGAGCCACAACCCCTCACTCGGCGCGCCAGTCCCTCCGATAGACTGCGTCCGCCGGGTA  
CCCGTATTTCCCAATAAAGCCTCTTGCTGTTTGATCCGAATCGTGGACTCGCTGATCCTTGGGAG  
GGTCTCCTCAGATTGATTGACTGCCACCTCGGGGTCTTTTCAATTTGGAGGTTCCACCGAGATTT  
GGAGACCCCTGCCAGGGACCACCGACCCCGCCGGGAGGTAAGCTGGCCAGCGGTGTTTTG  
TGTCTGTCTGTCTTTGTGCGTGTGTTGTGCGGCATCTAATGTTTGCCTGCGTCTGTACTAG  
TTAGCTTACTAGCTGTATCTGGCGGACCCGTGGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT  
CAACCCTGGGAGACTCCCAGGGACTTTGGGGCGTTTTTGTGGCCGACCTGAGGAAGGGAGT  
CGATGTGGAATCCGACCCCGTCAGGATATGTGGTCTGGTAGGAGACGAGAACC TAAAACAGTTT  
CCGCTCCGTCTGAATTTTTGCTTTGCTTTGGAACCGAAGCCGCGCTCTTGTCTGCTGCAGCG  
CTGCAGCATCGTTCTGTGTTGTCTCTGTCTGACTGTGTTTCTGTATTTGTCTGAAAATTAGGGC  
AGACTGTTACCACTCCCTTAAAGTTTGACCTTAGGTCAGTGGAAAGATGTGAGCGGATCGCTCAC  
AACCAGTCGGTAGATGTCAAGAAGAGACGTTGGGTTACCTTCTGCTCTGCAGAATGGCCAACCTT  
TAACGTCGGATGGCCGCGAGACGGCACCTTTAACCAGACCTCATCACCCAGGTTAAGATCAAG  
TCTTTTACCTGGCCCGCATGGACACCAGACCAGGTCCTTACATCGTGACCTGGGAAGCCTTG  
GCTTTTGACCCCCCTCCCTGGGTCAAGCCCTTTGTACACCCTAAGCCTCCGCCTCCTCTTCTCC  
ATCCGCCCCGTCTCTCCCCCTTGAACCTCCTCGTTCGACCCCGCTCGATCCTCCTTTATCCAG  
CCCTCACTCCTTCTCTAGGCGCCGGAATTAGATCTCTCGATAATAGGGGACCCGGATCCCCCGG  
TCCAACACCCGTTGGAAATCCCATGGGCACCATGGCCCCCTCGCTCAAAAATGCTTTCGCGTCTCG  
CAGACACTGCTCGGTAGTTTTCGGGATCAGCGTTTGTAGTAAGAGCCCGCTCTGAACCTCCGCG  
CCGCCCCGGCCAGTGGAAGACGCGCAGGCAAAACGCACCACGTGACGGAGCGTGACCGCGCGC  
CGAGCGCGCCAAAGTCCGGCAGGAAGAGGGCCTATTTCCCATGATTCCTTTCATATTTGCATAT  
ACGATACAAGGCTGTTAGAGAGATAATTAGAATTAATTTGACTGTAAACACAAAGATATTAGTAC  
AAAATACGTGACGTAGAAAGTAATAATTTCTTGGGTAGTTTGCAGTTTTTAAAATTTATGTTTTAA  
AATGGACTATCATATGCTTACCGTAACCTGAAAGTATTTGATTTCTTGGCTTTATATATCTTGT  
GGAAAGGACGAAACACCGTGTCTGCTTCCGCGAGCACATATACTAGTCTGACTAGGGATAACAGGGT  
AATTGTTTGAATGAGGCTTTCAGTACTTTACAGAATCGTTGCCTGCACATCTTGGAAACACTTGCT  
GGGATTACTTCTTTCAGGTTAACCCAACAGAAGGCTCGAGCAACCAGAATTCAGGGGCTACTTTA  
GGAGCAATTATCTTGTTTACTAAAACCTGAATACCTTGCTATCTCTTTGATACATTTTTACAAAGC  
TGAATTAAAATGGTATAAATTAATCACTTTTTTCAATTGGAAGACTAATGCGTTTAAACACGCG  
GCGACGCGTTTCGACGAATAAAAACCTGTGACGGAAGATCACTTCGAGAATAAAATAAATCCTGGT  
GTCCCTGTTGATACCGGAAGCCCTGGGCCAACCTTTGGCGAAAATGAGACGTTGATCGGCACGT  
AAGAGGTTCCAACCTTACCATAATGAAATAAGATCACTACCGGGCGTATTTTTTGTGTTGTCGA  
GATTTTTCAGGAGCTAAGGAAGCTAAAATGGAGAAAAAATCACTGGATATAACCACCGTTGATATA  
TCCAATGGCATCGTAAAGAACATTTTGGAGCATTTTCAGTCAGTTGCTCAATGTACCTATAACCA  
GACCGTTTCAGCTGGATATTACGGCCTTTTTTAAAGACCGTAAAGAAAAATAAGCACAAAGTTTTATC  
CGGCCTTTATTCACATTTCTTGCCCGCTGATGAATGCTCATCCGGAATTACGTATGGCAATGAAA  
GACGGTGTAGCTGGTGATATGGGATAGTGTTCACCTTGTACACCGTTTTTCCATGAGCAAACCTGA  
AACGTTTTTCATCGCTCTGGAGTGAATACCACGACGATTTCCGGCAGTTTCTACACATATATTTCG  
AAGATGTGGCGTGTACGGTGA AAAACCTGGCCTATTTCCCTAAAGGGTTTTATTGAGAATATGTTT  
TTGCTCTCAGCCAATCCCTGGGTGAGTTTACCAGTTTTGATTTAAACGTGGCCAATATGGACAA  
CTTCTTCGCCCCGTTTTTACCATGGGCAATATTTATACGCAAGGCGACAAGGTGCTGATGCCGC  
TGGCGATTCAGGTTTCATCATGCCGTTTTGTGATGGCTTCCATGTGCGCAGAATGCTTAATGAATTA  
CAACAGTACTGCGATGAGTGGCAGGGCGGGCGTAATTTTTTTAAGGCAGTTATTGGTGCCTTA

AACGCCTGGTTGCTACGCCTGAATAAGTGATAATAAGCGGATGAATGGCAGAAATTCGGATCTCG  
ACCGCGTTTTGGGCGGTGGCTCCCTGCCACGCGGCTCCGAACAGAAGCTGATCTCCGAAGAGGATC  
TGATTACCCTGTTATCCCTACCCTAAAATTCTACCGGGTAGGGGAGCGCTTTTCCCAAGGCAGT  
CTGGAGCATGCGCTTTAGCAGCCCCGCTGGGCACTTGGCGCTACACAAGTGGCCTCTGGCCTCGC  
ACACATTTCCACATCCACCGGTAGGCGCAACCGGCTCCGTTCTTTGGTGGCCCCCTTCGCGCCACC  
TTCTACTCCTCCCCTAGTCAGGAAGTTCCCCCGCCCCGAGCTCGCGTCGTGCAGGACGTGAC  
AAATGGAAGTAGCACGTCTCACTAGTCTCGTGCAGATGGACAGCACCGCTGAGCAATGGAAGCGG  
GTAGGCCTTTTGGGCGAGCGGCAATAGCAGCTTTGCTCCTTTCGCTTTCTGGGCTCAGAGGCTGGG  
AAGGGTGGGTCCGGGGCGGGCTCAGGGGCGGGCTCAGGGGCGGGGCGGGCGCCGAAGGTCTT  
CCGGAGGCCCGGCATTCTGCACGCTTCAAAGCGCACGTCTGCCGCGCTGTTCTCCTCTTCTCA  
TCTCCGGGCTTTTCGACCTGCAGCCCAAGCTTACCATGACCGAGTACAAGCCACGGTGCAGCTC  
GCCACCCGCGACGACGTCCCAGGGCCGTACGCACCCTCGCCGCGCGTTCGCCGACTACCCCGC  
CACGCGCCACACCGTCGATCCGGACCGCCACATCGAGCGGGTACCGAGCTGCAAGAACTCTTCC  
TCACGCGCGTCCGGCTCGACATCGGCAAGGTGTGGGTTCGCGGACGACGGCGCCGCGGTGGCGGT  
TGGACCACGCCGGAGAGCGTCAAGCGGGGGCGGTGTTCCGCGAGATCGGCCCGCGCATGGCCGA  
GTTGAGCGGTTCCCGGCTGGCCGCGCAGCAACAGATGGAAGGCCTCCTGGCGCCGACCCGGCCCA  
AGGAGCCCGCGTGGTTCCTGGCCACCGTCGGCGTCTCGCCCGACCACCAGGGCAAGGGTCTGGGC  
AGCGCGTCTGTGCTCCCCGAGTGGAGGCGGCCGAGCGCGCCGGGTGCCCGCTTCTGGAGAC  
TCCCGCGCCCCGCAACCTCCCCTTCTACGAGCGGCTCGGCTTACCCTCACCGCCGACGCTCGAGG  
GTCCCAGAAGGACCGCGACCTGGTGCATGACCCCGCAAGCCCGGTGCCTGACGCCGCCCCACGAC  
CCGACGCGCCGACCGAAAGGAGCGCACGCCCATGCATCGATAAAATAAAAGATTTTATTTAG  
TCTCCAGAAAAAGGGGGGAATGAAAAGCCCCACCTGTAGGTTTTGGCAAGCTAGAGAACCATCAGA  
TGTTTTCCAGGGTGCSCCAAGGACCTGAAATGACCCTGTGCCTTATTTGAACTAACCAATCAGTTC  
GCTTCTCGCTTCTGTTTCGCGCGCTTCTGCTCCCCGAGCTCAATAAAAGAGCCCAACCCCTCAC  
TCGGCGCGCCAGTCTCCGATAGACTGCGTTCGCCGGGTACCCGTGTATCCAATAAACCCCTCTTG  
CAGTTGCATCCGACTTGTGGTCTCGCTGTTCTTGGGAGGGTCTCCTCTGAGTGATTGACTACCC  
GTCAGCGGGGGTCTTTCATGGGTAAACGTTTTCTTGAAGTTGGAGAACAACATTCTGAGGGTAGGA  
GTCAATCGACGTGGGCCAATT  
CTGTCAGCCGTTAAGTGTTCCTGTGTCACTGAAAAATTGCTTTGAGAGGCTCTAAGGGCTTCTCAG  
TGCGTTACATCCCTGGCTTGTGTCCACAACCGTTAAACCTTAAAAGCTTTAAAAGCCTTATATA  
TTCTTTTTTTCTTATAAAACTTAAAACCTTAGAGGCTATTTAAGTTGCTGATTTATATTAATTT  
TATTGTTCAAACATGAGAGCTTAGTACGTGAAACATGAGAGCTTAGTACGTTAGCATTGAGAGCT  
TAGTACGTTAGCCATGAGGGTTTTAGTTCGTTAAACATGAGAGCTTAGTACGTTAAACATGAGAGC  
TTAGTACGTGAAACATGAGAGCTTAGTACGTAATCAACAGGTTGAACTGCTGATCAACAGATC  
CTCTACTAGAAAGGACGCACCGCTAGCAGCGCCCCTAGCGGTATCCTATAAAAAAACACACCG  
CGCCGCTAGCAGCACCCCTAATATAAAATAATGTTTTTTATAAAAAATAGTCAGTACCACCCCTAC  
AAAACGGTGTTCGGCGCTTGTGTAGCCGCGCCGACACCGCTTTTTTAAATATCATAAAGAGAGT  
AAGAGAAACTAATTTTTTCATAACACTCTATTTATAAAGAAAAATCAGCAAAAACTGTTTTTTGCG  
TGGGTGTGGTGTCTTTGGTGGTGAGAACCACCAACCTGTTGAGCCTTTTTTGTGGAGTGGGTAA  
ATTATACTAGCGCTTTCGAACCCCAGAGTCCCGCTCAGAAGAACTCGTCAAGAAGGCGATAGAA  
GGCGATGCGCTGCGAATCGGGAGCGCGATACCGTAAAGCACGAGGAAGCGGTGAGCCATTTCG  
CGCAAGCTCTTCAGCAATATCACGGGTAGCCAACGCTATGTCTGATAGCGGTCCGCCACACCC  
AGCCGGCCACAGTCGATGAATCCAGAAAAGCGGCCATTTTCCACCATGATATTGCGCAAGCAGGC  
ATCGCCATGTGTACGACGAGATCCTCGCCGTCGGGATGCGCGCCTTGAGCCTGCGCAACAGTT  
CGGCTGGCGCGACCCCTGATGCTCTTCGTCAGATCATCCTGATCGACAAGACCGCTTCCATC  
CGAGTACGTGCTCGCTCGATGCGATGTTTTCGCTTGGTGGTTCGAATGGGCAGGTAGCCGGATCAAG  
CGTATGCAGCCCGCCGATTCGATCAGCCATGATGGATACTTTCTCGGCAGGAGCAAGGTGAGATG  
ACAGGAGATCCTGCCCGGCACTTCGCCCAATAGCAGCCAGTCCCTTCCCGCTTCAGTGACAACG  
TCGAGCACAGCTGCGCAAGGAACGCCCGTCTGGCCAGCCACGATAGCCGCGCTGCCTCGTCTCTG  
CAGTTCATTACGGGCACCGGACAGGTTCGGTCTTGACAAAAAGAACCAGGGCGCCCTGCGCTGACA  
GCCGGAACACGGCGGCATCAGAGCAGCCGATTGTCTGTTGTGCCAGTCATAGCCGAATAGCCTC  
TCCACCCAAGCGGCCGGAGAACCTGCGTGAATCCATCTTGTTCATCATGCGAAACGATCCTCA  
TCCTGTCTCTTGATCAGATCT

**See Appendix 1 for restriction analysis**

Figure 5: pSM2 cloning region



See Appendix 2 for restriction analysis

## Appendix1

### Restriction analysis of pSM2 vector

#### SspI

1 CTTCCAACC TTACCAGAGG GCGCCCAGC TGTCCGAAAT ATTATAAATT  
GAAGGGTTGG AATGGTCTCC CGCGGGGTCG ACAGGCTTA TAATATTTAA

#### BstXI

51 ATCGCACACA TAAAACCAT GCTGTTGGTG TGTCTATTAA ATCGGCAACT  
TAGCGTGTGT ATTTTGGTA CGACAACCAC ACAGATAATT TAGCCGTTGA

#### PvuI

101 GTTGGGAAGG GCGATCGGTG CGGGCCTCTT CGCTATTACG CCAGCTGGCG  
CAACCTTCC CGCTAGCCAC GCCCGGAGAA GCGATAATGC GGTCGACCGC

151 AAAGGGGGAT GTGCTGCAAG GCGATTAAGT TGGGTAACGC CAGGGTTTTTC  
TTCCCCCTA CACGACGTTT CGCTAATTCA ACCCATTGCG GTCCCAAAG

201 CCAGTCACGA CGTTGTAAAA CGACGGCGCA AGGAATGGTG CATGCAAGGA  
GGTCAGTGCT GCAACATTTT GCTGCCGCGT TCCTTACCAC GTACGTTCT

251 GATGGCGCCC AACAGTCCCC CGGCCACGGG GCCTGCCACC ATACCCACGC  
CTACCGCGGG TTGTCAGGGG GCCGGTGCC CGGACGGTGG TATGGGTGCG

#### AfeI BspHI

301 CGAAACAAGC GCTCATGAGC CCGAAGTGGC GAGCCCGATC TTCCCCATCG  
GCTTTGTTTCG CGAGTACTCG GGCTTCACCG CTCGGGCTAG AAGGGGTAGC

#### SgrAI

351 GTGATGTCGG CGATATAGGC GCCAGCAACC GCACCTGTGG CGCCGGTGAT  
CACTACAGCC GCTATATCCG CGGTCGTTGG CGTGGACACC GCGGCCACTA

401 GCCGGCCACG ATGCGTCCGG CGTAGAGGCG ATTAGTCAA TTTGTAAAG  
CGGCCGGTGC TACGCAGGCC GCATCTCCGC TAATCAGGTT AAACAATTC

#### EcoRV

451 ACAGGATATC AGTGGTCCAG GCTCTAGTTT TGA CTCAACA ATATCACCAG  
TGTCTATAG TCACCAGGTC CGAGATCAA ACTGAGTTGT TATAGTGTC

501 CTGAAGCCTA TAGAGTACGA GCCATAGATA AAATAAAAGA TTTTATTTAG  
GACTTCGGAT ATCTCATGCT CGGTATCTAT TTTATTTTCT AAAATAAATC

551 TCTCCAGAAA AAGGGGGGAA TGAAAGACCC CACCTGTAGG TTTGGCAAGC  
AGAGGTCTTT TTCCCCCTT ACTTTCTGGG GTGGACATCC AAACCGTTCC

#### AfIII

601 TAGCTTAAGT AACGCCATTT TGCAAGGCAT GGAAAATACA TAACTGAGAA  
ATCGAATTCA TTGCGGTAAA ACGTTCCGTA CCTTTTATGT ATTGACTCTT

651 TAGAGAAGTT CAGATCAAGG TTAGGAACAG AGAGACAGCA GAATATGGGC  
ATCTCTCAA GTCTAGTCC AATCCTTGTC TCTCTGTCGT CTTATACCCG



GACCATCCTC TGCTCTTGA TTTTGTCAAG GCGGAGGCA GACTTAAAA

**AfeI**

1451 GCTTTCGGTT TGAACCGAA GCCGCGCCTC TTGTCTGCTG CAGCGCTGCA  
CGAAAGCCAA ACCTTGCTT CCGCAGCAG AACAGACGAC GTCGCGACGT

1501 GCATCGTTCT GTGTTGTCTC TGTCTGACTG TGTTTCTGTA TTTGTCTGAA  
CGTAGCAAGA CACAACAGAG ACAGACTGAC ACAAAGACAT AACAGACTT

**AflII Bsu36I**

1551 AATTAGGGCC AGACTGTTAC CACTCCCTTA AGTTTGACCT TAGGTCAGTG  
TTAATCCCGG TCTGACAATG GTGAGGGAAT TCAAAGTGA ATCCAGTGAC

1601 GAAAGATGTC GAGCGGATCG CTCACAACCA GTCGGTAGAT GTCAAGAAGA  
CTTTCTACAG CTCGCCTAGC GAGTGTGGT CAGCCATCTA CAGTTCTTCT

**BstEII**

1651 GACGTTGGGT TACCTTCTGC TCTGCAGAAT GGCCAACCTT TAACGTCGGA  
CTGCAACCCA ATGGAAGACG AGACGTCTTA CCGGTTGGAA ATTGCAGCCT

1701 TGGCCGCGAG ACGGCACCTT TAACCGAGAC CTCATCACCC AGGTTAAGAT  
ACGGCGCTC TGCCGTGGAA ATTGGCTCTG GAGTAGTGGG TCCAATTCTA

**SexAI**

1751 CAAGGTCTTT TCACCTGGCC CGCATGGACA CCCAGACCAG GTCCCCTACA  
GTTCCAGAAA AGTGGACCGG GCGTACCTGT GGGTCTGGTC CAGGGGATGT

1801 TCGTGACCTG GGAAGCCTTG GCTTTTGACC CCCCTCCCTG GGTCAAGCCC  
AGCACTGGAC CTTTCGGAAC CGAAAAGTGG GGGGAGGGAC CCAGTTCGGG

**BsrGI**

1851 TTTGTACACC CTAAGCCTCC GCCTCCTCTT CCTCCATCCG CCCCCTCTCT  
AAACATGTGG GATTCCGAGG CGGAGGAGAA GGAGGTAGGC GGGGCAGAGA

1901 CCCCCTTGAA CTCCTCGTT CGACCCCGCC TCGATCCTCC CTTTATCCAG  
GGGGAACTT GGAGGAGCAA GCTGGGGCGG AGCTAGGAGG GAAATAGGTC

**EcoNI**

**BglII**

1951 CCCTCACTCC TTCTCTAGGC GCCGGAATTA GATCTCTCGA TAATAGGGGA  
GGGAGTGAGG AAGAGATCCG CGGCCTTAAT CTAGAGAGCT ATTATCCCCT

**CspCI**

**BamHI**

**CspCI**

**XcmI**

2001 CCGGATCCCC CCGAGTCCAA CACCCGTGGG AATCCCATGG GCACCATGGC  
GGCCTAGGGG GGCTCAGGTT GTGGGCACCC TTAGGGTACC CGTGGTACCG

**BtsI**

2051 CCCTCGCTCC AAAAATGCTT TCGCGTCTCG CAGACACTGC TCGGTAGTTT  
GGGAGCGAGG TTTTACGAA AGCGCAGAGC GTCTGTGACG AGCCATCAA

2101 CGGGGATCAG CGTTTGAGTA AGAGCCCGCG TCTGAACCCT CCGCGCCGCC  
GCCCTAGTC GCAAACATCAT TCTCGGGCGC AGACTTGGGA GCGCGGGCGG

**PmlI**

**DrallI**

2151 CCGGCCAGT GGAAAGACGC GCAGGCAAAA CGCACCACGT GACGGAGCGT  
GGCCGGGTCA CCTTTCTGCG CGTCCGTTTT GCGTGGTGCA CTGCCTCGCA

2201 GACCGCGCGC CGAGCGCGCG CCAAGGTCGG GCAGGAAGAG GGCCTATTTT  
CTGGCGCGCG GCTCGCGCGC GGTTCAGCC CGTCCTTCTC CCGGATAAAG

2251 CCATGATTCC TTCATATTTG CATATACGAT ACAAGGCTGT TAGAGAGATA  
GGTACTAAGG AAGTATAAAC GTATATGCTA TGTTCCGACA ATCTCTCTAT

**AseI**

2301 ATTAGAATTA ATTTGACTGT AAACACAAAG ATATTAGTAC AAAATACGTG  
TAATCTTAAT TAAACTGACA TTTGTGTTTC TATAATCATG TTTTATGCAC

2351 ACGTAGAAAG TAATAATTTT TTGGGTAGTT TGCAGTTTTT AAAATTATGT  
TGCATCTTTC ATTATTAAG AACCCATCAA ACGTCAAAAA TTTTAATACA

**NdeI**

2401 TTTAAAATGG ACTATCATAT GCTTACCGTA ACTTGAAAGT ATTTGATTT  
AAATTTTACC TGATAGTATA CGAATGGCAT TGAACCTTCA TAAAGCTAAA

2451 CTTGGCTTTA TATATCTTGT GGAAAGGACG AAACACCGTG CTCGCTTCGG  
GAACCGAAAT ATATAGAACA CCTTCTCTGC TTTGTGGCAC GAGCGAAGCC

**Sall**

**HincII**

**AccI**

2501 CAGCACATAT ACTAGTCGAC TAGGGATAAC AGGGTAATTG TTTGAATGAG  
GTCGTGTATA TGATCAGCTG ATCCCTATTG TCCCATTAAC AAACCTACTC

**Scal**

2551 GCTTCAGTAC TTTACAGAAT CGTTGCCTGC ACATCTTGGA AACACTTGCT  
CGAAGTCATG AAATGTCTTA GCAACGGACG TGTAGAACCT TTGTGAACGA

**HpaI**

**HincII**

**XhoI**

**PspXI**

**EcoRI**

2601 GGGATTACTT CTTCAAGTTA ACCCAACAGA AGGCTCGAGC AACCGAATT  
CCCTAATGAA GAAGTCCAAT TGGGTTGTCT TCCGAGCTCG TTGGTCTTAA

2651 CAAGGGGCTA CTTTAGGAGC AATTATCTTG TTTACTAAAA CTGAATACCT  
GTTCCCGAT GAAATCCTCG TTAATAGAAC AAATGATTTT GACTTATGGA

2701 TGCTATCTCT TTGATACATT TTTACAAAGC TGAATTAATAA TGGTATAAAT  
ACGATAGAGA AACTATGTAA AAATGTTTCG ACTTAATTTT ACCATATTA

**MluI**

**MfeI**

**BbsI**

**PmeI**

**AflIII**

2751 TAAATCACTT TTTTCAATTG GAAGACTAAT GCGTTTAAAC ACGCGGCGAC  
ATTTAGTGAA AAAAGTTAAC CTTCTGATTA CGCAAATTTG TGCGCCGCTG

2801 GCGTTCGACC GAATAAAACC TGTGACGGAA GATCACTTCG CAGAATAAAT  
CGCAAGCTGG CTTATTTTGG AACTGCCTT CTAGTGAAGC GTCTTATTTA

**PfIMI**

2851 AAATCCTGGT GTCCCTGTTG ATACCGGGAA GCCCTGGGCC AACTTTTGGC  
TTTAGGACCA CAGGGACAAC TATGGCCCTT CGGGACCCGG TTGAAAACCG

**MslI**

2901 GAAAATGAGA CGTTGATCGG CACGTAAGAG GTTCCAACCT TCACCATAAT  
CTTTACTCT GCAACTAGCC GTGCATTCTC CAAGGTTGAA AGTGGTATTA

2951 GAAATAAGAT CACTACCGGG CGTATTTTTT GAGTTGTCGA GATTTTCAGG  
CTTTATTCTA GTGATGGCCC GCATAAAAAA CTCAACAGCT CAAAAGTCC

3001 AGCTAAGGAA GCTAAAATGG AGAAAAAAT CACTGGATAT ACCACCGTTG  
TCGATTCTT CGATTTTACC TCTTTTTTA GTGACCTATA TGGTGGCAAC

3051 ATATATCCCA ATGGCATCGT AAAGAACATT TTGAGGCATT TCAGTCAGTT  
TATATAGGGT TACCGTAGCA TTTCTTGTA AACTCCGTAA AGTCAGTCAA

3101 GCTCAATGTA CCTATAACCA GACCGTTCAG CTGGATATTA CGGCCTTTTT  
CGAGTTACAT GGATATTGGT CTGGCAAGTC GACCTATAAT GCCGGAAAAA

3151 AAAGACCGTA AAGAAAAATA AGCACAAGTT TTATCCGGCC TTTATTCA  
TTTCTGGCAT TTCTTTTTAT TCGTGTTCAA AATAGGCCGG AAATAAGTGT

**SnaBI**

3201 TTCTTGCCCG CCTGATGAAT GCTCATCCGG AATTACGTAT GGCAATGAAA  
AAGAACGGGC GGACTACTTA CGAGTAGGCC TTAATGCATA CCGTTACTTT

3251 GACGGTGAGC TGGTGATATG GGATAGTGTT CACCCTTGTT ACACCGTTTT  
CTGCCACTCG ACCACTATAC CCTATCACAA GTGGGAACAA TGTGGCAAAA

**AclI**

3301 CCATGAGCAA ACTGAAACGT TTTCATCGCT CTGGAGTGAA TACCACGACG  
GGTACTCGTT TGACTTTGCA AAAGTAGCGA GACCTCACTT ATGGTGCTGC

3351 ATTTCCGGCA GTTTCTACAC ATATATTCGC AAGATGTGGC GTGTTACGGT  
TAAAGCCGT CAAAGATGTG TATATAAGCG TTCTACACCG CACAATGCCA

3401 GAAAACCTGG CCTATTTCCC TAAAGGGTTT ATTGAGAATA TGTTTTTCGT  
CTTTTGACC GGATAAAGGG ATTTCCCAA TAACTCTTAT AAAAAAGCA

**PfIMI**

3451 CTCAGCCAAT CCCTGGGTGA GTTTCACCAG TTTTGATTTA AACGTGGCCA  
GAGTCGGTTA GGGACCCACT CAAAGTGGTC AAAACTAAAT TTGCACCGGT

**SspI**

3501 ATATGGACAA CTTCTTCGCC CCCGTTTTCA CCATGGGCAA ATATTATACG  
TATACCTGTT GAAGAAGCGG GGGCAAAAGT GGTACCCGTT TATAATATGC

3551 CAAGGCGACA AGGTGCTGAT GCCGCTGGCG ATTCAGGTTT ATCATGCCGT  
GTTCCGCTGT TCCACGACTA CGGCGACCGC TAAGTCCAAG TAGTACGGCA

**Scal**

3601 TTGTGATGGC TTCCATGTCG GCAGAATGCT TAATGAATTA CAACAGTACT  
AACACTACCG AAGGTACAGC CGTCTTACGA ATTAATAAT GTTGTATGCA

3651 GCGATGAGTG GCAGGGCGGG GCGTAATTTT TTTAAGGCAG TTATTGGTGC  
CGCTACTCAC CGTCCCGCCC CGCATTAATA AAATTCCGTC AATAACCACG

3701 CCTTAAACGC CTGGTTGCTA CGCCTGAATA AGTGATAATA AGCGGATGAA  
GGAATTTGCG GACCAACGAT GCGGACTTAT TCACTATTAT TCGCCTACTT

3751 TGGCAGAAAT TCGGATCTCG ACCGCGTTTG GGCGGTGGCT CCCTGCCACG  
ACCGTCTTTA AGCCTAGAGC TGGCGCAAAC CCGCCACCGA GGGACGGTGC

3801 CGGCTCCGAA CAGAAGCTGA TCTCCGAAGA GGATCTGATT ACCCTGTTAT  
GCCGAGGCTT GTCTTCGACT AGAGGCTTCT CTTAGACTAA TGGGACAATA

3851 CCCTACCCTA AAATTCTACC GGGTAGGGGA GGCGCTTTTC CCAAGGCAGT  
GGATGGGAT TTTAAGATGG CCCATCCCCT CCGCGAAAAG GGTTCCGTC

3901 CTGGAGCATG CGCTTTAGCA GCCCCGCTGG GCACTTGGCG CTACACAAGT  
GACCTCGTAC GCGAAATCGT CGGGGCGACC CGTGAACCGC GATGTGTTCA

**AgeI**

3951 GGCCTCTGGC CTCGCACACA TTCCACATCC ACCGGTAGGC GCCAACCGGC  
CCGGAGACCG GAGCGTGTGT AAGGTGTAGG TGGCCATCCG CGGTTGGCCG

4001 TCCGTTCTTT GGTGGCCCCT TCGCGCCACC TTCTACTCCT CCCCTAGTCA  
AGGCAAGAAA CCACCGGGGA AGCGCGGTGG AAGATGAGGA GGGGATCAGT

4051 GGAAGTTCCC CCCC GCCCCG CAGCTCGCGT CGTGCAGGAC GTGACAAATG  
CCTTCAAGGG GGGGCGGGG GTCGAGCGCA GCACGTCCTG CACTGTTTAC

**BssSI**

**B1pI**

4101 GAAGTAGCAC GTCTACTAG TCTCGTGCAG ATGGACAGCA CCGCTGAGCA  
CTTCATCGTG CAGAGTGATC AGAGCACGTC TACCTGTCGT GGCGACTCGT

**StuI**

4151 ATGGAAGCGG GTAGGCCTTT GGGGCAGCGG CCAATAGCAG CTTTGCTCCT  
TACCTTCGCC CATCCGGAAA CCCCCTCGCC GGTATCGTC GAAACGAGGA

4201 TCGCTTTCTG GGCTCAGAGG CTGGGAAGGG GTGGGTCCGG GGGCGGGCTC  
AGCGAAAGAC CCGAGTCTCC GACCCTTCCC CACCCAGGCC CCCGCCGAG

4251 AGGGGCGGGC TCAGGGGCGG GCGGGCGCC CGAAGGTCCT CCGGAGGCC  
TCCC CGCCG AGTCCC CGCC CCGCCCGGG GCTTCCAGGA GGCCTCCGGG

4301 GGCATTCTGC ACGTTCAAAG AGCGCACGTC TGCCGCGCTG TTCTCCTCTT  
CCGTAAGACG TCGAAGTTT TCGCGTGCAG ACGGCGCGAC AAGAGGAGAA

**HindIII**

4351 CCTCATCTCC GGGCCTTTTCG ACCTGCAGCC CAAGCTTACC ATGACCGAGT  
GGAGTAGAGG CCCGGAAGC TGGACGTCGG GTTCGAATGG TACTGGCTCA

**BsiWI**

4401 ACAAGCCAC GGTGCGCCTC GCCACCCGCG ACGACGTCCC CAGGGCCGTA  
TGTTCCGGTG CCACGCGGAG CGGTGGGCGC TGCTGCAGGG GTCCC GGCAT

4451 CGCACCTCG CCGCCGCGTT CGCCGACTAC CCCGCCACGC GCCACACCGT  
GCGTGGGAGC GCGGCGCAA GCGGCTGATG GGGCGGTGCG CCGTGTGGCA

Ppil

RsrII            BstEII            Ppil

4501 CGATCCGGAC CGCCACATCG AGCGGGTCAC CGAGCTGCAA GAACTCTTCC  
GCTAGGCCTG GCGGTGTAGC TCGCCAGTG GCTCGACGTT CTTGAGAAGG

4551 TCACGCGCGT CGGGCTCGAC ATCGGCAAGG TGTGGGTCGC GGACGACGGC  
AGTGCGCGCA GCCCGAGCTG TAGCCGTTCC ACACCCAGCG CCTGCTGCC

SacII

4601 GCCGCGGTGG CGGTCTGGAC CACGCCGGAG AGCGTCGAAG CGGGGGCGGT  
CGGCGCCACC GCCAGACCTG GTGCGGCCTC TCGCAGCTTC GCCCCCGCCA

4651 GTTCGCCGAG ATCGGCCCGC GCATGGCCGA GTTGAGCGGT TCCCGGCTGG  
CAAGCGGCTC TAGCCGGGCG CGTACCGGCT CAACTCGCCA AGGGCCGACC

BsaXI

StuI                    BsaXI

4701 CCGCGCAGCA ACAGATGGAA GGCCTCCTGG CGCCGCACCG GCCCAAGGAG  
GGCGCGTCGT TGTCTACCTT CCGGAGGACC GCGGCGTGGC CGGGTTCTCT

4751 CCCGCGTGGT TCCTGGCCAC CGTCGGCGTC TCGCCCGACC ACCAGGGCAA  
GGGCGACCA AGGACCGGTG GCAGCCGCAG AGCGGGCTGG TGTCCCGTT

Bcgl

Bcgl

4801 GGGTCTGGGC AGCGCCGTCG TGCTCCCCGG AGTGGAGGCG GCCGAGCGCG  
CCCAGACCCG TCGCGGCAGC ACGAGGGGCC TCACCTCCGC CGGCTCGCGC

PfoI

4851 CCGGGGTGCC CGCCTTCTG GAGACCTCCG CGCCCCGCAA CCTCCCCTTC  
GGCCCCACGG GCGGAAGGAC CTCTGGAGGC GCGGGGCGTT GGAGGGGAAG

4901 TACGAGCGGC TCGGCTTAC CGTCACCGCC GACGTGAGG TGCCCGAAGG  
ATGCTCGCCG AGCCGAAGTG GCAGTGGCGG CTGCAGCTCC ACGGGCTTCC

SexAI

DraIII

4951 ACCGCGCACC TGGTGCATGA CCCGCAAGCC CCGTGCCTGA CGCCCGCCCC  
TGGCGGTGG ACCACGTACT GGGCGTTCGG GCCACGGACT GCGGGCGGGG

NsiI ClaI

BfrBI

5001 ACGACCCGCA GCGCCCGACC GAAAGGAGCG CACGACCCCA TGCATCGATA  
TGCTGGGCGT CGCGGGCTGG CTTTCTCGC GTGCTGGGGT ACGTAGCTAT

5051 AAATAAAAGA TTTTATTTAG TCTCCAGAAA AAGGGGGGAA TGAAAGACCC  
TTTATTTTCT AAAATAAATC AGAGGTCTTT TTCCCCCTT ACTTTCTGGG

5101 CACCTGTAGG TTTGGCAAGC TAGAGAACCA TCAGATGTTT CCAGGGTGCC

---

GTGGACATCC AAACCGTTCG ATCTCTTGGT AGTCTACAAA GGTCCCACGG

5151 CCAAGGACCT GAAATGACCC TGTGCCTTAT TTGAACTAAC CAATCAGTTC  
GGTTCCTGGA CTTTACTGGG ACACGGAATA AACTTGATTG GTTAGTCAAG

Sacl  
EcoICRI

5201 GCTTCTCGCT TCTGTTGCGC CGCTTCTGCT CCCCAGGCTC AATAAAAGAG  
CGAAGAGCGA AGACAAGCGC GCGAAGACGA GGGGCTCGAG TTATTTTCTC

Xmal  
SmaI

AscI

5251 CCCACAACCC CTCACTCGGC GCGCCAGTCC TCCGATAGAC TCGTTCGCCC  
GGGTGTTGGG GAGTGAGCCG CGCGGTCAGG AGGCTATCTG ACGCAGCGGG

BaeI  
BaeI  
KpnI  
BaeI  
BaeI  
Acc65I

5301 GGGTACCCGT GTATCCAATA AACCTCTTG CAGTTGCATC CGACTTGTGG  
CCCATGGGCA CATAGGTTAT TTGGGAGAAC GTCAACGTAG GCTGAACACC

5351 TCTCGCTGTT CTTGGGAGG GTCTCCTCTG AGTGATTGAC TACCCGTCAG  
AGAGCGACAA GGAACCCTCC CAGAGGAGAC TACTAACTG ATGGGCAGTC

5401 CGGGGGTCTT TCATGGGTAA CAGTTTCTTG AAGTTGGAGA ACAACATTCT  
GCCCCAGAA AGTACCCATT GTCAAAGAAC TTCAACCTCT TGTTGTAAGA

5451 GAGGGTAGGA GTCGAATCGA GAGAGAGAGA GAGAGAGAGA GAGAGAGAGA  
CTCCATCCT CAGCTTAGCT CTCTCTCTCT CTCTCTCTCT CTCTCTCTCT

PspOMI  
ApaI

5501 GAGAGAGAGA GACGTGGGCC CAATTCTGTC AGCCGTTAAG TGTTCTGTG  
CTCTCTCTCT CTGCACCCGG GTTAAGACAG TCGGCAATTC ACAAGGACAC

5551 TCACTGAAAA TTGCTTTGAG AGGCTCTAAG GGCTTCTCAG TCGTTACAT  
AGTGACTTTT AACGAACTC TCCGAGATTC CCGAAGAGTC ACGCAATGTA

HindIII

5601 CCCTGGCTTG TTGTCCACAA CCGTTAAACC TTAAAAGCTT TAAAAGCCTT  
GGGACCGAAC AACAGGTGTT GGCAATTTGG AATTTTCGAA ATTTTCGGAA

5651 ATATATTCTT TTTTTCTTA TAAAACCTAA AACCTTAGAG GCTATTTAAG  
TATATAAGAA AAAAAAGAAT ATTTTGAATT TTGGAATCTC CGATAAATTC

AseI

5701 TTGCTGATTT ATATTAATTT TATTGTTCAA ACATGAGAGC TTAGTACGTG  
AACGACTAAA TATAATTAATA ATAACAAGTT TGTACTCTCG AATCATGCAC

5751 AAACATGAGA GCTTAGTACG TTAGCCATGA GAGCTTAGTA CGTTAGCCAT

---

TTGTACTCT CGAATCATGC AATCGGTACT CTCGAATCAT GCAATCGGTA

5801 GAGGGTTT TAG TTCGTAAAC ATGAGAGCTT AGTACGTAA ACATGAGAGC  
CTCCCAAATC AAGCAATTG TACTCTCGAA TCATGCAATT TGTACTCTCG

**SnaBI**

5851 TTAGTACGTG AACATGAGA GCTTAGTACG TACTATCAAC AGGTTGAACT  
AATCATGCAC TTTGTACTCT CGAATCATGC ATGATAGTTG TCCAACCTGA

**BclI**

5901 GCTGATCAAC AGATCCTCTA CACTAGAAGG GACGCACCGC TAGCAGCGCC  
CGACTAGTTG TCTAGGAGAT GTGATCTTCC CTGCGTGGCG ATCGTCGCGG

5951 CCTAGCGGTA TCCTATAAAA AAACACACCG CGCCGCTAGC AGCACCCCTA  
GGATCGCCAT AGGATATTTT TTTGTGTGGC GCGGCGATCG TCGTGGGGAT

6001 ATATAAATA ATGTTTTTTA TAAAAATAGT CAGTACCACC CCTACAAAAC  
TATATTTTAT TACAAAAAAT ATTTTTATCA GTCATGGTGG GGATGTTTTG

6051 GGTGTCGGCG CGTTGTTGTA GCCGCGCCGA CACCGCTTTT TAAATATCA  
CCACAGCCGC GCAACAACAT CGGCGCGGCT GTGGCGAAAA AATTTATAGT

6101 TAAAGAGAGT AAGAGAACT AATTTTTCAT AACACTCTAT TTATAAAGAA  
ATTTCTCTCA TTCTCTTGA TAAAAAAGTA TTGTGAGATA AATATTTCTT

6151 AAATCAGCAA AAACCTGTTT TTGCGTGGGG TGTGGTGCTT TTGGTGGTGA  
TTTAGTCGTT TTTGAACAAA AACGCACCCC ACACCACGAA AACCACCACT

6201 GAACCACCAA CCTGTTGAGC CTTTTTGTGG AGTGGGTAA ATTATACTAG  
CTTGGTGGTT GGACAACCTG GAAAAACACC TCACCCAATT TAATATGATC

**BstBI BpII**

6251 CGCGTTTCGA ACCCCAGAGT CCCGCTCAGA AGAACTCGTC AAGAAGGCGA  
GCGCAAAGCT TGGGGTCTCA GGGCGAGTCT TCTTGAGCAG TTCTTCCGCT

**BssSI**

6301 TAGAAGGCGA TCGCTGCGA ATCGGGAGCG GCGATACCGT AAAGCACGAG  
ATCTTCCGCT ACGCGACGCT TAGCCCTCGC CGCTATGGCA TTTCTGTGCTC

**SapI**

6351 GAAGCGGTCA GCCCATTGCG CGCCAAGCTC TTCAGCAATA TCACGGGTAG  
CTTCGCCAGT CGGGTAAGCG GCGGTTGAG AAGTCGTTAT AGTGCCCATC

**RsrII**

6401 CCAACGCTAT GTCCTGATAG CGGTCCGCCA CACCCAGCCG GCCACAGTCG  
GGTTGCGATA CAGGACTATC GCCAGGCGGT GTGGGTCGGC CGGTGTCAGC

**MsiI**

6451 ATGAATCCAG AAAAGCGGCC ATTTTCCACC ATGATATTCG GCAAGCAGGC  
TACTTAGGTC TTTTCGCCGG TAAAAGGTGG TACTATAAGC CGTTCGTCCG

**BpuEI**

6501 ATCGCCATGT GTCACGACGA GATCCTCGCC GTCGGGCATG CGCGCCTTGA  
TAGCGGTACA CAGTGCTGCT CTAGGAGCGG CAGCCCCTAC GCGCGGAACT

**SapI**

6551 GCCTGGCGAA CAGTTCGGCT GGC GCGAGCC CCTGATGCTC TTCGTCCAGA  
CGGACCGCTT GTCAAGCCGA CCGCGCTCGG GGACTACGAG AAGCAGGTCT

**FalI**

6601 TCATCCTGAT CGACAAGACC GGCTTCCATC CGAGTACGTG CTCGCTCGAT  
AGTAGGACTA GCTGTTCTGG CCGAAGGTAG GCTCATGCAC GAGCGAGCTA

6651 GCGATGTTTC GCTTGGTGGT CGAATGGGCA GGTAGCCGGA TCAAGCGTAT  
CGCTACAAAG CGAACCACCA GCTTACCCGT CCATCGGCCT AGTTCGCATA

6701 GCAGCCGCCG CATTGCATCA GCCATGATGG ATACTTTCTC GGCAGGAGCA  
CGTCGGCGGC GTAACGTAGT CGGTACTACC TATGAAAGAG CCGTCCTCGT

6751 AGGTGAGATG ACAGGAGATC CTGCCCCGGC ACTTCGCCCA ATAGCAGCCA  
TCCACTCTAC TGTCTCTAG GACGGGGCCG TGAAGCGGGT TATCGTCGGT

**FspI**

6801 GTCCCTTCCC GCTTCAGTGA CAACGTCGAG CACAGCTGCG CAAGGAACGC  
CAGGGAAGGG CGAAGTCACT GTTGACGCTC GTGTCGACGC GTTCCTTGCG

6851 CCGTCGTGGC CAGCCACGAT AGCCGCGCTG CCTCGTCCTG CAGTTCATTC  
GGCAGCACCG GTCGGTGCTA TCGGCGCGAC GGAGCAGGAC GTCAAGTAAG

**DrdI**

6901 AGGGCACCGG ACAGGTCGGT CTTGACAAAA AGAACCGGGC GCCCTGCGC  
TCCCGTGGCC TGTCAGCCA GAACTGTTTT TCTTGCCCG CGGGGACGCG

6951 TGACAGCCGG AACACGGCGG CATCAGAGCA GCCGATTGTC TGTTGTGCC  
ACTGTCGGCC TTGTGCCGCC GTAGTCTCGT CGGCTAACAG ACAACACGGG

7001 AGTCATAGCC GAATAGCCTC TCCACCCAAG CGGCCGAGGA ACCTGCGTGC  
TCAGTATCGG CTTATCGGAG AGGTGGGTTT GCCGGCCTCT TGGACGCACG

**BsaBI**

**BclI**

7051 AATCCATCTT GTTCAATCAT GC GAAACGAT CCTCATCCTG TCTCTTGATC  
TTAGGTAGAA CAAGTTAGTA CGCTTTGCTA GGAGTAGGAC AGAGA ACTAG

**BglII**

7101 AGATCT

## Appendix 2: Restriction Analysis of pSM2 cloning region

EcoO109I

EarI Sau96I

1 GCAGGAAGAG GGCCTATTC CCATGATTCC TTCATATTTG CATATACGAT  
CGTCCTTCTC CCGGATAAAG GGTACTAAGG AAGTATAAAC GTATATGCTA

AseI

51 ACAAGGCTGT TAGAGAGATA ATTAGAATTA ATTTGACTGT AAACACAAAG  
TGTTCCGACA ATCTCTCTAT TAATCTTAAT TAAACTGACA TTTGTGTTTC

Tsp45I

101 ATATTAGTAC AAAATACGTG ACGTAGAAAG TAATAATTC TGGGGTAGTT  
TATAATCATG TTTTATGCAC TGCATCTTTC ATTATTAAAG AACCCATCAA

NdeI

151 TGCAGTTTTT AAAATTATGT TTTAAAATGG ACTATCATAT GCTTACCGTA  
ACGTCAAAAA TTTTAATACA AAATTTTACC TGATAGTATA CGAATGGCAT

201 ACTTCAAAGT ATTTGATTT CTTGGCTTTA TATATCTTGT GGAAAGGACG  
TGAACCTTCA TAAAGCTAAA GAACCGAAAT ATATAGAACA CCTTTCCTGC

Sall

HincII

AccI

Bsp1286I

TseI

Bfal

BsiHKAI

BbvI

SpeI

Bfal

251 AACACCGTG CTCGCTTCGG CAGCACATAT ACTAGTCGAC TAGGGATAAC  
TTTGTGGCAC GAGCGAAGCC GTCGTGTATA TGATCAGCTG ATCCCTATTG

Scal

AcuI

BsgI

301 AGGGTAATTG TTTGAATGAG GCTTCAGTAC TTTACAGAAT CGTTGCCTGC  
TCCCATTAAC AAACCTACTC CGAAGTCATG AAATGTCTTA GCAACGGACG

HpaI

BseYI

AcuI HincII

351 ACATCTTGA AACACTTGCT GGGATTACTT CTTCAAGTTA ACCCAACAGA  
TGTAGAACCT TTGTGAACGA CCCTAATGAA GAAGTCCAAT TGGGTTGTCT

XhoI

SmlI

AvaI

EcoRI

PspXI

ApoI

401 AGGCTCGAGC AACCGAATT CAAGGGGCTA CTTTAGGAGC AATTATCTTG  
TCCGAGCTCG TTGGTCTTAA GTTCCCCGAT GAAATCCTCG TTAATAGAAC

451 TTTACTAAAA CTGAATACCT TGCTATCTCT TTGATACATT TTTACAAAGC  
AAATGATTTT GACTTATGGA ACGATAGAGA AACTATGTAA AAATGTTTCG

MfeI BbsI

501 TGAATTAATA TGGTATAAAT TAAATCACTT TTTTCAATTG GAAGACTAAT  
ACTTAATTTT ACCATATTTA ATTTAGTGAA AAAAGTTAAC CTTCTGATTA

MluI  
AflIII  
HgaI TaqII TspGWI  
PmeI Hpy99I BsiEI Tsp45I  
551 GCGTTTAAAC ACGCGGCGAC GCGTTCGACC GAATAAAACC TGTGACGGAA  
CGCAAATTTG TGCGCCGCTG CGCAAGCTGG CTTATTTTGG ACACTGCCTT

BsmFI  
BsiFI  
BsiFI NciI  
601 GATCACTTCG CAGAATAAAT AAATCCTGGT GTCCCTGTTG ATACCGGGAA  
CTAGTGAAGC GTCTTATTTA TTTAGGACCA CAGGGACAAC TATGGCCCTT

Sau96I BsmBI  
PaeI PfiMI BsmAI  
651 GCCCTGGGCC AACTTTTGGC GAAAATGAGA CGTTGATCGG CACGTAAGAG  
CGGGACCCGG TTGAAAACCG CTTTACTCT GCAACTAGCC GTGCATTCTC

MmeI MspI NciI  
701 GTTCCAATT TCACCATAAT GAAATAAGAT CACTACCGGG CGTATTTTTT  
CAAGGTTGAA AGTGGTATTA CTTTATTCTA GTGATGGCCC GCATAAAAAA

DdeI  
Bpu10I TspRI  
751 GAGTTGTCGA GATTTTCAGG AGCTAAGGAA GCTAAAATGG AGAAAAAAT  
CTCAACAGCT CTAAGTCC TCGATTCTT CGATTTTACC TCTTTTTTTA  
M E K K I Frame 1

BsrI  
TspRI SfaNI  
801 CACTGGATAT ACCACCGTTG ATATATCCCA ATGGCATCGT AAAGAACATT  
GTGACCTATA TGGTGGCAAC TATATAGGGT TACCGTAGCA TTTCTTGTA  
T G Y T T V D I S Q W H R K E H F Frame 1

PvuII  
MspA1I  
851 TTGAGGCATT TCAGTCAGTT GCTCAATGTA CCTATAACCA GACCGTTCAG  
AACTCCGTAA AGTCAGTCAA CGAGTTACAT GGATATTGGT CTGGCAAGTC  
E A F Q S V A Q C T Y N Q T V Q Frame 1

BceAI  
901 CTGGATATTA CGGCCTTTTT AAAGACCGTA AAGAAAAATA AGCACAAGTT  
GACCTATAAT GCCGGAAAAA TTTCTGGCAT TTCTTTTTAT TCGTGTTCAA  
L D I T A F L K T V K K N K H K F Frame 1

BspEI  
BsaWI  
FokI  
FauI BsmI BstF5I  
951 TTATCCGGCC TTTATTCACA TTCTTGCCCG CCTGATGAAT GCTCATCCGG  
AATAGGCCGG AAATAAGTGT AAGAACGGGC GGACTACTTA CGAGTAGGCC  
Y P A F I H I L A R L M N A H P E Frame 1

SnaBI BsrDI  
1001 AATTACGTAT GGCAATGAAA GACGGTGAGC TGGTGATATG GGATAGTGTT  
TTAATGCATA CCGTTACTTT CTGCCACTCG ACCACTATAC CCTATCACAA  
L R M A M K D G E L V I W D S V Frame 1

AclI BtgZI  
1051 CACCCTTGTT ACACCGTTTT CCATGAGCAA ACTGAAACGT TTTCATCGCT  
GTGGGAACAA TGTGGCAAAA GGTACTCGTT TGACTTTGCA AAAGTAGCGA  
H P C Y T V F H E Q T E T F S S L Frame 1

BpmI Hpy99I  
1101 CTGGAGTGAA TACCACGACG ATTTCCGGCA GTTCTACAC ATATATTCGC  
GACCTCACTT ATGGTGCTGC TAAAGGCCGT CAAAGATGTG TATATAAGCG  
W S E Y H D D F R Q F L H I Y S Q Frame 1

1151 AAGATGTGGC GTGTTACGGT GAAAACCTGG CCTATTTCCC TAAAGGGTTT  
TTCTACACCG CACAATGCCA CTTTTGGACC GGATAAAGGG ATTTCCCAA  
D V A C Y G E N L A Y F P K G F Frame 1

DdeI  
BspCNI  
BseMII  
BsmAI TaqII  
BsmBI PfiMI PstI BsrI  
1201 ATTGAGAATA TGTTTTTCGT CTCAGCCAAT CCCTGGGTGA GTTTCACCAG  
TAACTCTTAT ACAAAAAGCA GAGTCGGTTA GGGACCCACT CAAAGTGGTC  
I E N M F F V S A N P W V S F T S Frame 1

MscI  
EaeI  
1251 TTTTGATTTA AACGTGGCCA ATATGGACAA CTTCTTCGCC CCCGTTTTCA  
AAAATAAAT TTGCACCGGT TATACCTGTT GAAGAAGCGG GGGCAAAAGT  
F D L N V A N M D N F F A P V F T Frame 1

StyI  
NcoI MspA1I  
BtgI SspI SfaNI  
1301 CCATGGGCAA ATATTATACG CAAGGCGACA AGGTGCTGAT GCCGCTGGCG  
GGTACCCGTT TATAATATGC GTTCCGCTGT TCCACGACTA CGGCGACCGC  
M G K Y Y T Q G D K V L M P L A Frame 1

BceAI BclI BsmI  
1351 ATTCAGGTTT ATCATGCCGT TTGTGATGGC TTCCATGTCTG GCAGAATGCT  
TAAGTCCAAG TAGTACGGCA AACACTACCG AAGGTACAGC CGTCTTACGA  
I Q V H H A V C D G F H V G R M L Frame 1

Scal BtgZI Faul  
1401 TAATGAATTA CAACAGTACT GCGATGAGTG GCAGGGCGGG GCGTAATTTT  
ATTACTTAAT GTTGTGATGA CGTACTCAC CGTCCC GCCC CGCATTAAAA  
N E L Q Q Y C D E W Q G G A \* Frame 1

Bsp1286I  
Bme1580I  
BanI

1451 TTTAAGGCAG TTATTGGTGC CCTTAAACGC CTGGTTGCTA CGCCTGAATA  
AAATTCCGTC AATAACCACG GGAATTTGCG GACCAACGAT GCGGACTTAT

FokI BstYI  
BstF5I ApeI AlwI BsiEI

1501 AGTGATAATA AGCGGATGAA TGGCAGAAAT TCGGATCTCG ACCGCGTTTG  
TCACTATTAT TCGCCTACTT ACCGTCTTTA AGCCTAGAGC TGGCGCAAAC

AlwNI EarI

1551 GGCGGTGGCT CCCTGCCACG CGGCTCCGAA CAGAAGCTGA TCTCCGAAGA  
CCGCCACCGA GGGACGGTGC GCCGAGGCTT GTCTTCGACT AGAGGCTTCT

BstYI  
AlwI

1601 GGATCTGATT  
CCTAGACTAA

## Useful References:

Carmell, M.A., Zhang, L., Conklin, D.S., Hannon, G.J., Rosenquist, T.A., **Germline transmission of RNAi in mice**. Nat Struct Bio 10, 91-92. (2003).

Conklin, D., Hannon, G., Bernstein, E., Caudy, A., Paddison, P., **Short Hairpin RNAs (shRNAs) induce sequence-specific silencing in Mammalian Cells**. Genes & Development. 16:948-958. (2002).

Hannon, G., Caudy, A., Paddison, P., **Stable Suppression of Gene Expression by RNAi in Mammalian Cells**. PNAS 99(3): 1443-1448 (2003).

Hemann, M.T., Friedman, J.S., Zilfou, J.T., Hernando, E., Paddison, P., Cordon-Cardo, C., Hannon, G.J., Lowe, S.W. **An epi-allelic series of p53 hypomorphs created by stable RNAi produces distinct tumor phenotypes in vivo**. Nat Genet 33, 396-400. (2003).

McCaffrey A.P., Meuse L., Pham T.T., Conklin D.S., Hannon G.J., Kay M.A. **RNA interference in adult mice**. Nature (418) 38-39. (2002).

Paddison, P.J., Silva, J.M., Conklin, D.S., Schlabach, M., Li, M., Aruleba, S., Balija, V., O'Shaughnessy, A., Gnoj, L., Scobie, K. Chang, K., Westbrook, T., Sachidanandam, R., McCombie W.R., Elledge, S.J., & Hannon, G.J. **A resource for large-scale RNAi based screens in mammals**. Nature, March 25; 427-431. (2004)

Paddison, P.J., Caudy, A.A., Bernstein, E., Hannon, G.J. & Conklin, D.S. **Short hairpin RNAs (shRNAs) induce sequence-specific silencing in mammalian cells**. Genes Dev 16,948-58. (2002).

Taira, K., Miyagishi, M. **U6 promoter-driven siRNAs with four uridine 3' overhangs efficiently suppress targeted gene expression in mammalian cells**. Nature Biotechnology 19, 497-500. (2002).

Zeng, Yan., Cullen, B. **Sequence requirements for micro RNA processing and function in human cells**. RNA 9, 112-123. (2003).

**Limited Use License:**

*This product is covered by several patent applications owned by Cold Spring Harbor Laboratory. The purchase of this product conveys to the buyer the limited, non-exclusive, non-transferable right (without the right to resell, repackage, or further sublicense) under these patent rights to perform the RNAi knockdown methods using the RNAi-inducing vectors claimed in those patent applications for research purposes solely in conjunction with this product. No other license is granted to the buyer whether expressly, by implication, by estoppel or otherwise. In particular, the purchase of this product does not include nor carry any right or license to use, develop, or otherwise exploit this product commercially, and no rights are conveyed to the buyer to use the product or components of the product for any other purposes, including without limitation, provision of services to a third party, generation of commercial databases, or clinical diagnostics or therapeutics.*

*This product is sold pursuant to a license from CSHL, and CSHL reserves all other rights under these patent rights. For information on purchasing a license to the patent rights for uses other than in conjunction with this product or to use this product for purposes other than research, please contact the CSHL Office of Technology Transfer at, 516-367-8312.*

---

Phone: 1-888-412-2225

FAX: 1-256-704-4849

[info@openbiosystems.com](mailto:info@openbiosystems.com)

V010304

For Research Use Only