Bisulfite sequencing

1. Design PCR primers for amplification of the CpG-enriched region
    a. For the primer design, use Zymo Research online tool: http://www.zymoresearch.com/tools/bisulfite-primer-seeker
    b. For the promoter, 200-1000bp upstream of the +1 site is commonly used. It's better to contain >10 CpG islands within the amplicon.
    c. Remember to keep the original sequence for the analysis.
2. Purify genomic DNA with phenol/chloroform method or kits
3. Perform bisulfite reaction using EZ DNA Methylation-Gold™ Kit (Zymo Research, #D5005) following the manual
    a. 0.5-2μg DNA can be used for the reaction. In general, 0.5μg is sufficient for the test of 3-4 loci. Use 2μg if you are going test more loci.
    b. Bisulfite-treated DNAs are eluted with 10μl elution buffer twice for total 20 μl
4. Perform PCR reaction using the Jumpstart RedTaq (Sigma, #P2893-100RXN)
    a. 1-2 μl eluate can be used as the template. Start with 1 μl l then go to 2 μl l if it doesn't work.
    b. Reaction volume can be 25 or 50μl. 25μl should be sufficient in most cases. If you don't get bright band, increase the volume.
5. Gel purify the PCR product and clone it to the pGEM-T (or other TA cloning) vector (no need to do A-tailing).
6. Sequencing >10 colonies containing the insert (regular Sanger method is fine and should be tried first. Sometimes specific condition for the secondary structure will be needed)
7. Analyze the results using QUMA: http://quma.cdb.riken.jp/
    a. This website needs two files: one target sequence file containing the original (non-bisulfite-converted) sequence and bisulfite sequence file containing complied sequences of colonies got from step 6.
    b. Remember to remove the primer sequences since they are not gone through the bisulfite conversion.
    c. The incomplete sequence can also be used.

Ex.

The target file contains:

```
>Oct4promoter
GCTGTCTTGTCCTGGCCTTGGACATGGGCTGAAATACTGGGTTCACCCATATCTAGGACTCTAGACGG
GTGGGTAAGCAAGAACTGAGGAGTGGCCCCAGAAATAATTGGCACACGAACATTCAATGGATGTTTTA
GGCTCTCCAGAGGATGGCTGAGTGGGCTGTAAGGACAGGCCGAGAGGGTGCAGTGCCAACAGGCTTTG
TGGTGCGATGGGGCATCCGAGCAACTGGTTTGTGAGGTGTCCGGTGACCCAAGGCAGGGGTGAGAGGA
CCTTGAAGGTTGAAAATGAAGGCCTCCTGGGGTCCCGTCCTAAGGGTTGTCCTGTCCAGACGTCCCCA
```

ACCTCCGTCTGGAAGACACAGGCAGATAGCGCTCGCCTCAGTTTCTCCCACCCCCACAGCTCTGCTCC
TCCACCCACCCAGGGGGCGGGGCCAGAGGTCAAGGCTAGAGGGTGGGA

And the bisulfite sequence file contains:

>Oct4wt_1
TGCTCCCGGCCGCCATGGCGGCCGCGGGAATTCGATTGCGTTATCGAGGTCGTTGTTTTGTTTTGGTT
TTGGATATGGGTTGAAATATTGGGTTTATTTATATTTAGGATTTTAGATGGGTGGGTAAGTAAGAATT
GAGGAGTGGTTTTAGAAATAATTGGTATATGAATACTTAATGGATGTTTTAGGTTTTTTAGGGGATGG
TTGAGTGGGTTGTAAGGATAGGTTGAGAGGGTGTAGTGTTAATAGGTTTTGTGGTGTGATGGGGTATT
TGAGTAATTGGTTTGTGAGGTGTTTGGTGATTTAAGGTAGGGGTGAGAGGATTTTGAAGGTTGAAAAT
GAAGGTTTTTTGGGGTTTTGTTTTAAGGGTTGTTTTGTTTAGATGTTTTTAATTTTTGTTTGGAAGAT
ATAGGTAGATAGTGTTTGTTTTAGTTTTTTTTATTTTTATAGTTTTGTTTTTTTATTTATTTAGGGGG
TGGGGTTAGAGGTTAAGGTTAGAGGGTGGGATCGGAAGAGCACAATCACTAGTGAATTCGCGGCCGCC
TGCAGGTCGACCATATGGGAGAGCTCCCAACGCGTTGGATGCATAGCTTGAGTATTCTATAGTGTCAC
CTAAATAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCCGCTCACAATTC
CNCACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGNCTAATGAGTGAGCTAACTCACA
TTTAATTGCGTTTGCGCTCACTGNCCNGCTTTCCAGTCGGGNAAACCTTGTCGTGCCAGCTGCATTAA
TNGANNCGGNCNA

>Oct4wt_2
GNNGCCNTGGCGGCCGCGGGAATTCGATTGCGTTATCGAGGTCGTTGTTTTGTTTTGGTTTTGGATAT
GGGTTGAAATATTGGGTTTATTTATATTTAGGATTTTAGATGGGTGGGTAAGTAAGAATTGAGGAGTG
GTTTTAGAAATAATTGGTATATGAATATTTAATGGATGTTTTAGGTTTTTTAGAGGATGGTTGAGTGG
GTTGTAAGGATAGGTTGAGAGGGTGTAGTGTTAATAGGTTTTGTGGTGTGATGGGGTATTTGAGTAAT
TGGTTTGTGAGGTGTTTGGTGATTTAAGGTAGGGGTGAGAGGATTTTGAAGGTTGAAAATGAAGGTTT
TTTGGGGTTTTGTTTTAAGGGTTGTTTTGTTTAGATGTTTTTAATCTTTGTTTGGAAGATATAGGTAG
ATAGTGTTTGTTTTAGTTTTTTTTATTTTTATAGTTTTGTTTTTTTATTTATTTAGGGGGTGGGGTTA
GAGGTTAAGATTAGAGGGTGGGATCGGAAGAGCACAATCACTAGTGAATTCGCGGCCGCCTGCAGGTC
GACCATATGGGAGAGCTCCCAACGCGTTGGATGCATAGCTTGAGTATTCTATAGTGTCACCTAAATAG
CTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAAATTGTTATCCGCTCACAATTCCACACNAA
CATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGNCTAATGAGTGAGCTAACTCACATTAATTG
CGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCNGGCCA
ACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCT
CGGTCGTTCGGCTGCGGCGAGCGGTATCAGCNNANNTCANGGCGGTAATACGGTTATCCACAGNATCN
GNATAACGCAGGAANANCATTGTGAGCAAAGGCCNGCAAAGGCANNACCGTAAAAGGCCGCGTTGCTG
GCGTTTTCATNGCTCCGCCCCTGACGAGCATCANAAAATCGNNNTNCAGTCANANGTGGNNAACCCGA
NGGANTANNNNNNNNCCANGNG

>Oct4wt_3
ATGCTNNGGCCGCCATGGCGGCCGCGGGAATTCGATTGCGTTATCGAGGTCGTTGTTTTGTTTTGGTT
TTGGATATGGGTTGAAATATTGGGTTTATTTATATTTAGGATTTTAGATGGGTGGGTAAGTAAGAATT
GAGGAGTGGTTTTAGAAATAATTGGTATATGAATATTTAATGGATGTTTTAGGTTTTTTAGAGGATGG
TTGAGTGGGTTGTAAGGATAGGTTGAGAGGGTGTAGTGTTAATAGGTTTTGTGGTGTGATGGGGTATT
TGAGTAATTGGTTTGTGAGGTGTTTGGTGATTTAAGGTAGGGGTGAGAGGATTTTGAAGGTTGAAAAT
GAAGGTTTTTTGGGGTTTTGTTTTAAGGGTTGTTTTGTTTAGATGTTTTTAATTTTTGTTTGGAAGAT
ATAGGTAGATAGTGTTTGTTTTAGTTTTTTTTATTTTTATAGTTTTGTTTTTTTATTTATTTAGGGGG
TGGGGTTAGAGGTTAAGGTTAGAGGGTGGGATCGGAAGAGCACAATCACTAGTGAATTCGCGGCCGCC
TGCAGGTCGACCATATGGGAGAGCTCCCAACGCGTTGGATGCATAGCTTGAGTATTCTATAGTGTCAC
CTAAATAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCC
ACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACAT
TAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATC
GGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTNCCGCTTCTCGCTCACTGACTCGCTG

CGCTCGGTCGTTCGGCTGCGGCGNAGCGGTATCAGCTCACTCAAGGCGGTAATACGGTTATCCACAGA
ATCAGGGNANACGCANGAAAGAACATGTGAGCAAAGGCCAGCAAANGCCAGGAACCGTAAAAGGCGCG
TTGNTGGCGTTTTCNNNNGCTCCGCCCNCCTGANNANCATCANAAANTNNNCGCTCAAGTCANNGTGG
NGAAACCNGACAGAANNNNNAANNNNCAGGNNTTCCCCTGANNCTNGCTCGTGNNCTCNNNNNTCCGA
NNNNCNNNACCGNNNACTNGNNCNNNTNNNNNTTNN

etc.