



DISH: Construction of Conditional Targeting Vector

Sections:

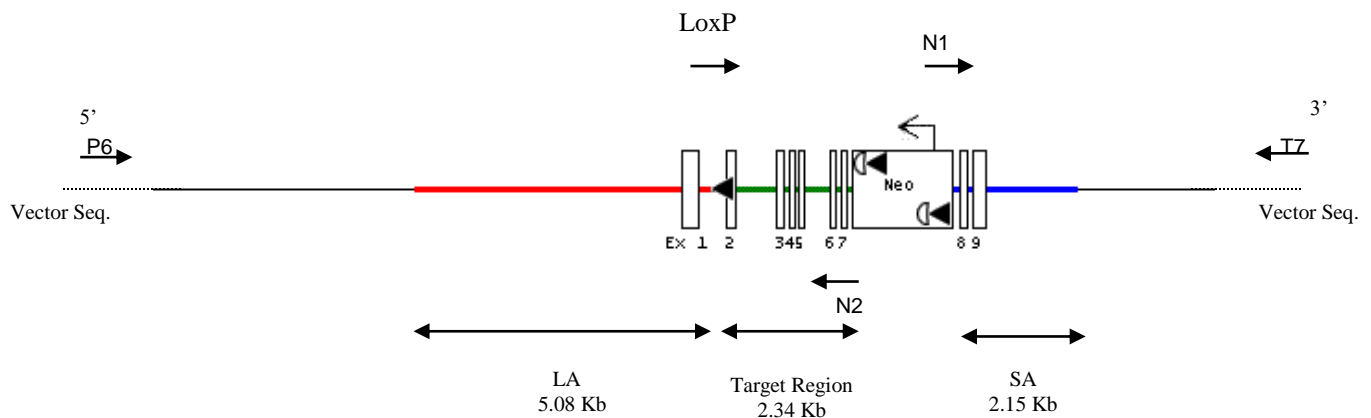
1. Vector Design Outline
2. Sequence Data Analysis
3. Cassette Sequences

1. Vector Design Outline

A 9.57 Kb region used to construct the targeting vector was first subcloned from a positively identified C57BL/6 BAC clone (RP23: 104G23). The region was designed such that the short homology arm (SA) extends about 2.15 Kb 3' to exon 7. The long homology arm (LA) ends 5' to exon 2, and is 5.08 Kb long. The loxP/FRT flanked Neo cassette is inserted 121 bp downstream of exon 7. The single loxP site, containing engineered Apal, BamHI and Mfe I sites for southern blot analysis, is inserted 176 bp upstream of exon 2. The target region is 2.34 Kb and includes exons 2-7.

The targeting vector is confirmed by restriction analysis after each modification step. P6 and T7 primers anneal to the backbone vector sequence and read into the 5' and 3' ends of the BAC sub-clone. N1 and N2 primers anneal to the 5' and 3' ends of the LoxP/FRT Neo cassette and sequence the SA and LA, respectively.

BAC Subclone and confirmation



PCR primers used for sequencing:

Primer P6: 5' - GAG TGC ACC ATA TGG ACA TAT TGT C -3'
Primer T7: 5' - CGA TAA GCC AGG TTA ACC TGC ATT A -3'
Primer N1: 5' - TGC GAG GCC AGA GGC CAC TTG TGT AGC -3'
Primer N2: 5' - TTC CTC GTG CTT TAC GGT ATC G -3'
Primer loxP: 5' - GTA AAC TAG AGT TCT GCG ATG GGT GG -3'



Backbone Vector Information

The BAC was sub cloned into a ~2.4kb backbone vector (pSP72, Promega) containing an ampicillin selection cassette for retransformation of the construct prior to electroporation. A pGK-gb2 loxP/FRT Neo cassette was inserted into the gene as described in the project schematic. The targeting construct can be linearized using Asc I prior to electroporation into ES cells.

The total size of the targeting construct (including vector backbone and Neo cassette) is 13.67 Kb.



2. Sequence Data Analysis

Sequencing Data from BAC Sub Clone and Targeting Construct

P6 sequencing data aligned with genomic sequence

```
Query 7444 CAGGGCTGGGGTGCCCTGCTGGTCTGTGTTTGCATGTGCTCAAGGGCCATGGACACCTGTG 7503
          |||
Sbjct 126 CAGGGCTGGGGTGCCCTGCTGGTCTGTGTTTGCATGTGCTCAAGGGCCATGGACACCTGTG 185

Query 7504 TCTGATATAACTCAGTGAGCTGTTGGTGGGCTGTGTGGCATGCAAACCTGACTCAGAGA 7563
          |||
Sbjct 186 TCTGATATAACTCAGTGAGCTGTTGGTGGGCTGTGTGGCATGCAAACCTGACTCAGAGA 245

Query 7564 GCAGACACATGTTCCAGCCACCCTGTAGGTATGTGGCAGGTGGCCATAGTGCAGGTCTCT 7623
          |||
Sbjct 246 GCAGACACATGTTCCAGCCACCCTGTAGGTATGTGGCAGGTGGCCATAGTGCAGGTCTCT 305

Query 7624 TCATGCAGGGCTCTAGATGTACAGATATGAGCAGCTGTGACTGGAAATGGCCATAGGAGG 7683
          |||
Sbjct 306 TCATGCAGGGCTCTAGATGTACAGATATGAGCAGCTGTGACTGGAAATGGCCATAGGAGG 365

Query 7684 AGAATGACAGGTATGCACTGGTCCCCAGACACCTGTGGTGGCCCTCACCTTCCCATTCCAG 7743
          |||
Sbjct 366 AGAATGACAGGTATGCACTGGTCCCCAGACACCTGTGGTGGCCCTCACCTTCCCATTCCAG 425

Query 7744 CACACAAATTGCAACAGAGAAGCTGTCATCTTTATTGAGGACCTTCTGTGGTATCCAGA 7803
          |||
Sbjct 426 CACACAAATTGCAACARAGAARCAAGTTCATCTTTATTGAGGACCTTCTGTGGTATCCAGA 485

Query 7804 GTAATAAAGGCAGCAGCATTATCATGGGTTTCATAGGATAGCACAGGGTTCTTTTCCTTG 7863
          |||
Sbjct 486 GTAATAAAGGCAGCAGCATTATCATGGGTTTCATAGGATAGCACAGGGTTCTTTTCCTTG 545

Query 7864 AAAGGGTTGGTGGCAGCTGGGATGCCACAAGGAAGGGGTCGCTCTTGGCCTGTTTCAGTG 7923
          |||
Sbjct 546 AAAGGGTTGGTGGCAGCTGGGATGCCACAAGGAAGGGGTCGCTCTTGGCCTGTTTCAGTG 605

Query 7924 CAGAAGTGCAGCAGATCTGTAGCTGCCTTGGATACCTGCAGACCCC-ATCATGTCAGCCA 7982
          |||
Sbjct 606 CAGAAGTGCAGCAGATCTGTAGCTGCCTTGGATACCTGCAGACCCCATCATGTCAGCCA 665

Query 7983 CATCTGGGTATCCACCTCCTCCCTACACAGCTGTGCCACCCCCAGGCTGTCCCAACCCT 8042
          |||
Sbjct 666 CATCTGGGTATCCACCTCCTCCCTACACAGCTGTGCCACCCCCAGGCTGTCCCAACCCT 725

Query 8043 CACCTTTATGCGGTTGATGCCTGCTTCCATCCGAGTTGCTCCACTGCCGACGGGCTG 8102
          |||
Sbjct 726 CACCTTTATGCGGTTGATGCCTGCTTCCATCCGAGTTGCTCCACTGCCGACGGGCTG 785

Query 8103 TCCGATGTCA-CTGCCAGTGGCCACCTTGCTAGA-CATCTGGGAGAAGG-TGCGC-AGGG 8158
          |||
Sbjct 786 TCCGATGTCACTGCCAGTGGCCACCTTGCTTARAACATCTGGGAGAAGGKTGSGCCAGGG 845

Query 8159 CCGTGCAAGCCTGAGACTCC-AAGCTTTT-GCAGC-TCCTTTTAGTCTCTG 8206
          |||
Sbjct 846 CCGTGCAAGCYTGA-ACTCCTAAGCTTTTGCARCCTC-TTTAAG-TCCTG 893
```

Query: Genomic Sequence from ENSEMBL Database
Sbjct : Sequencing data from Targeting Vector



N1 sequencing data aligned with genomic sequence

```
Query  14862  GGATGGTTTGTGGCAGCCTCAGGCACTGATGGGCCCTTCCTAAAAGCTCTCTAGGCCTTC 14921
|||||
Sbjct   218      GGATGGTTTGTGGCAGCCTCAGGCACTGATGGGCCCTTCCTAAAAGCTCTCTAGGCCTTC 277

Query  14922  CTCTGCATCATACCCAGCCTTCTACGGCTTCTGACGAGGGCTTCCTCTACCCACAGAGGA 14981
|||||
Sbjct   278      CTCTGCATCATACCCAGCCTTCTACGGCTTCTGACGAGGGCTTCCTCTACCCACAGAGGA 337

Query  14982  ATGGAGCAGACTACGCTGTTTATATCAACACAGCCAGGAGTTTGATGGCTCAGACTCCG 15041
|||||
Sbjct   338      ATGGAGCAGACTACGCTGTTTATATCAACACAGCCAGGAGTTTGATGGCTCAGACTCCG 397

Query  15042  GAGCCCGGCCAGATGAGGCTGTCTCTTGGGGCAAGATCCGGATGGACGCACAGCCAGTAA 15101
|||||
Sbjct   398      GAGCCCGGCCAGATGAGGCTGTCTCTTGGGGCAAGATCCGGATGGACGCACAGCCAGTAA 457

Query  15102  AGGTAGATGGTGGCAGGGCAGGGTGCAGGGTCCGTGGGCTGCCCATGGCTCCAGGGCTCA 15161
|||||
Sbjct   458      AGGTAGATGGTGGCAGGGCAGGGTGCAGGGTCCGTGGGCTGCCCATGGCTCCAGGGCTCA 517

Query  15162  CTGAGGTCTTCTGCAGGTCTATGCTGATGCTTCTCTGGTTTTCCCTTGCTGGTGGCTGA 15221
|||||
Sbjct   518      CTGAGGTCTTCTGCAGGTCTATGCTGATGCTTCTCTGGTTTTCCCTTGCTGGTGGCTGA 577

Query  15222  GACATTCGCCCCAAAAGGCAGATGCCTTCAGAGCTGAGAAGAATGAAGACTAAGAAGATTG 15281
|||||
Sbjct   578      GACATTCGCCCCAAAAGGCAGATGCCTTCAGAGCTGAGAAGAATGAAGACTAAGAAGATTG 637

Query  15282  GTAAAGACTGAGGCTTTTTGCCACACCTTTATTTATTACTTTACATGCCAGCCCCCTCCCT 15341
|||||
Sbjct   638      GTAAAGACTGAGGCTTTTTGCCACACCTTTATTTATTACTTTACATGCCAGCCCCCTCCCT 697

Query  15342  AGGCCCACTCCTTGGTCAGCAGCATCTCTAGAATAAATGGCCCTTT-GGTGGGTTTTCT 15400
|||||
Sbjct   698      AGGCCCACTCCTTGGTCAGCAGCATCTCTAGAATAAATGGCCCTTTGGTGGGTTTTCT 757

Query  15401  GAGTCAGTCTGCATCCTTGGTGTGTCAGGTCGGTAAGCCTCAGCCTGCTCCACCCTCTGCCT 15460
|| |
Sbjct   758      GAK-CAGTCTGCATCCTTGGTGTGTCAGGTCGGTAAGCCTCAGCCTGCTCCACC-TCTGC-T 814

Query  15461  CATAGGTCTCTTCTCGCCAGTACTGGATGCTGCCTCCCATGGCAGTCAGTAGGCAGGGTT 15520
|||||
Sbjct   815      CATAGGTCTCTTCTCGCCAGTACTGGATGCTGCCTCC-ATGGCAGTCAK-AGGCAGGGTT 872

Query  15521  CCGTGGGATGGTAAGCCAGTGA-CTGTACCACATTAGAACCCACAGGCAGGGCCAGTG-C 15578
|||||
Sbjct   873      CCGTGGGATGGTAAGC-AGTGAMCTGTACCA-ATAA-AACCCACAG-CAGGGC-AGTGTC 927

Query  15579  CAAGGCACC-CTAAGGAATGAGTGGTGGGGA 15608
|| |
Sbjct   928      CA-G-CAACTCTAAG-AAKRAGTG-TGGGGA 954
```

Query: Genomic Sequence from ENSEMBL Database
Sbjct : Sequencing data from Targeting Vector



N1 Raw sequencing data (LoxP highlighted in yellow, FRT underline)

ACGCCAAGTGCCCAGCGGGGCTGCTAAAGCGCAGTCCAGACTGCCTTGGGAAAAGCGCCTCCCCTACCCGG
TAGAATGAAGTTCCTATACTTTCTAGAGAATAGGAACTTCGTTCGAACATAACTTCGTATAGCATACATTA
TACGAAGTTATGGTACCTGCAGAATTCATGCATAAGCTTGGATCCGTTCTTCGGACGCCTCGTCAACACCG
TACGGGATGGTTTGTGGCAGCCTCAGGCACCTGATGGGCCCTTCCTAAAAGCTCTCTAGGCCTTCCTCTGCA
TCATAACCCAGCCTTCTACGGCTTCTGACGAGGGCTTCCTCTACCCACAGAGGAATGGAGCAGACTACGCTG
TTTATATCAACACAGCCCAGGAGTTTGATGGCTCAGACTCCGGAGCCCGCCAGATGAGGCTGTCTCTTGG
GGCAAGATCCGGATGGACGCACAGCCAGTAAAGGTAGATGGTGGCAGGGCAGGGTGCAGGGTCCGTGGGCT
GCCCATGGCTCCAGGGCTCACTGAGGTCTTCTGCAGGTCTATGCTGATGCTTCTCTGGTTTTCCCCTTGCT
GGTGGCTGAGACATTCGCCAAAAGGCAGATGCCTTCAGAGCTGAGAAGAATGAAGACTAAGAAGATTGGT
AAAGACTGAGGCTTTTTGCCACACCYTTATTTATTACTTTACATGCCAGCCCCTCCCTAGGCCACTCCTT
GGTCAGCAGCATCTCTAGAATAAATGGGCCTTTTGGTTGGGTTTTCTGAKCAGTCTGCATCCTTGGTGTCA
GGTCCGTAAGCCTCAGCCTGCTCCACCTCTGCTCATAGGTCTCTTCTCGCCAGTACTGGATGCTGCCTCCA
TGGCAGTCAKAGGCAGGGTTCCGTGGGATGGTAAGCAGTGAMCTGTACCAATAAAAACCCACAGCAGGGCAG
TGTCCAGCAACTCTAAGAAKRAAGTGTGGGGATCATTMTA



N2 Raw sequencing data (LoxP highlighted in yellow, FRT underlined)

ACGGCTCGATTTCGMAGCGCATCGCCTTCTATCGCTTCTTGACGAGTWCTTCTGAGCGGGACTCTGGGGTTTC
GAATAAAGACCGACCAAGCGACGTCTGAGAGCTCCCTGATATCAGATCTACATAACTTCGTATAATGTATG
CTATACGAAGTTATCCGCGGGAAGTTCCCTATTCTCTAGAAAGTATAGGAACTTCGCTAGACTAGTACGCGT
GTACACTTAAGCCGGCGTACGGAATTCGCTAGCAGGAGACCACACCTACCTCTGCTCCCACTACAGCACTA
TTCCATGCCTCCGATCCATTCCCCTTACCCACAGCCACCATATAACACAGTGGGGACGGTCACAGGCCA
CTGCACTCACCATGAGGTTAGCGTTGGCAATGTGGTGCTTGACCACACCTCCACCCAGGATGATCATCCCA
GAGCGCTTGGCGAAAATGGCCTGCGTGTGATGAGTCGCAGGTCTGTGAGAGGGCAGAGGTGAGGCCTTGA
CCTCGGGGGCTTTTCTCCCTCCCTCCCTCCCGGGCTTAGCACCTCACCTTCAACAATGTCCAGGACCAAGC
CTGGGTTTTTATAGGAATGGAAGAAGATCATGTACCCAGTGAAGCATCTGTGAGTGCTGGACTCAGCACA
GGGATGTGGTTCTGGTGGAGACAAGACMAGTCAGCAGCTACAGCTTACACCCAGGAAAACCCTGACCCC
CAACCCAGTGGCGTCCCTCCCAACCTGTGTAGCCACTCAATTTCTCTGCTGGACATGACCCTCTGTTCTAC
ATTGTGGAGGCAGAGTGARGAAAACCACACATGGCTTGAAGGTGCTGGCMCMYTGAGCMCTGGCTTGTAC
TTCCAGACTCTTGTGCTGCACTGTCCCTTGGGGCACAGTACCATGGGTCCAGCCCCACTKGACTTCTCWAG
GGT

LoxP Raw sequencing data (LoxP highlighted in yellow, Mfe I highlighted in green, BamHI highlighted in teal, Apa I highlighted in pink). The LoxP cassette is 77 bps.

AAGTTGGAGTTAAAATCMAAAGGGCTTGAGGAGGAGCTTACGCTAGGGGAGCCCGGGGTTTCGATTTTCCCG
GTGGAAGGAGACTTGAAGATACTGGGTGGTTCTCAAAGGATCCATCACGCGTATAACTTCGTATAGCATA
CATTATACGAAGTTATGCCACTAGAGGATCCCCGCAATTGGGCCCATCAAGTGACAGACCTGGACCTGGAA
TGGGAGGGTTAGGGTCTGCCCAGGGTGTGCTTTGGTTTTGTAGTAGTCTCCCTGAGTGCAATCAACCTC
TGGAATCCAGGCCAGGTAGGGTACCTCAACTGAGGTTTGGAGTCTTAGTGCAACTTTCCTGTGCCCTTATG
ACCAGATTGAGAAAAAACTGGAGCCACTGGCTGTAGATGAAGATCATCACGCAGACCTGACTCAGAGCCGC
CGCCCACTTACAGGCTGCACCAATTTCTTGGGCTATACTTCCAACCTCATCAGTTCAGGCATCCGGGAGAC
CATTTCGATACCTCGTGCAGCACAACATGGTGGGTCCACGTTAGCTCTGGAACCCAGCCTCCTTGGGCCAG
TAGGAAAAAAAAAAAAAAAAAGGGTTATTTTTTTCAGTTAAGAGAAAATGAAGAGGCTCAGAGATTATTTGCGGTCTC
GTGAAAGATTTGGGTTTGTCTTTTTAACACCACATGGTGACTCAGAACCCGTCTGTAATGGGATCTTTTGAT
GCCATTACAGATCTGGTGTGTCTGAAGAAGAGGGACAGTGTATTACATATATTAATAACATCTTTTAGTC
TTTTTTGTTTTTGTGTTTGGAGACAGGTTTTCTCTGTATAGCCTGGCTGTCMTCCTGAACTCACTYCTGT
AGAACAAGGCTGGCCTTGAACCTMAAATTCTGCCTGCCTCCTACTCCCAGGTGCTTGGGAATAAAGGCATGC
CCTCACACCCGCCCAAATGATTTAAAAAAGTGT

NB/ Raw sequencing data is aligned with the genomic sequence predicted from the ENSEMBL database.



3. Cassette Sequences

Backbone Vector Sequence

3' end of BAC subclone joins here

CGATGATATCAGATCTGCCGGTCTCCCTATAGTGAGTCGTATTAATTTTCGATAAGCCAGGTTAACCTGCAT
TAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGA
CTCGCTGCGCTCGGTCGTTCCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCA
CAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAG
GCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCA
GAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTTCCCTTGGAAAGCTCCCTCGTGCGCTCTC
CTGTTCCGACCCTGCCGTTACC GGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCAT
AGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCCGCTCCAAGCTGGGCTGTGTGCACGAACCCCC
CGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTAT
CGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTG
AAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTAC
CTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTTGTTT
GCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGAC
GCTCAGTGAACGAAAACCTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGAT
CCTTTTAAATTAATAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACC
AATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTTCGTTTCATCCATAGTTGCCTGACTCCCC
GTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCC
ACGCTCACCGGCTCCAGATTTATCAGCAATAAACAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCCTG
CAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAAT
AGTTTGGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTACGCTCGTCTGTTGGTATGGCTTCATT
CAGCTCCGGTTCCTAACGATCAAGGCGAGTTACATGATCCCCATGTTGTGCAAAAAGCGGTTAGCTCCT
TCGGTCCCTCCGATCGTTGTGAGAAGTAAGTTGGCCGAGTGTATCACTCATGGTTATGGCAGCACTGCAT
AATTCTCTTACTGTGATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAAGTCAACCAAGTCAATCTG
AGAATAGTGTATGCGGCGACCGAGTTGCTCTTCCCGGCGTCAATACGGGATAATACCGGCCACATAGCA
GAACCTTAAAGTGTCTCATCATTGGAACCGTTCTTCCGGGCGAAAACCTCTCAAGGATCTTACCGCTGTTG
AGATCCAGTTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTACCAGCGTTTC
TGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATAC
TCATACTCTTCTTTTTCAATATTATTGAAGCATTATCAGGGTTATTGTCTCATGAGCGGATACATATTT
GAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCGAAAAGTGCCACCTGACGTCTA
AGAAACCATTATTATCATGACATTAACCTATAAAAATAGGCGTATCACGAGGCCCTTTTCGTCTCGCGGTT
TCGGTGTGACGGTGAACCTCTGACACATGCAGCTCCCGGAGACGGTACAGCTTGTCTGTAAGCGGAT
GCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGTGGCGGGTGTGCGGGCTGGCTTAACTATGC
GGCATCAGAGCAGATTGTAAGTGCACCATATGGACATATTGTCGTTAGAACGCGGCTACAATTAAT
ACATAACCTTATGTATCATAACATACGATTTAGGTGACACTATACCTGCAGGCGCGCCATTTAATGCGG
CCGCTTAATTAACCTGA

5' end of BAC subclone joins here



LoxP/FRT PGK-gb2-Neo Cassette (3' – 5' orientation)

LoxP sites are highlighted in yellow

FRT sites are ***Bold/Italics***

N1 and N2 sequencing primers are underlined

ATG and **STOP** codons are indicated

TAGCGAATTCCGTACGCCGGCTTAAGTGTACACGCGTACTAGTCTAGC**GAAGTTCCTATACTTTCTAGAGA**
ATAGGAACTTCCCGCGG**ATAACTTCGTATAGCATAACATTATACGAAGTTAT**GTAGATCTGATATCAGGGAG
CTCTCAGACGTCGCTTGGTCGGTCTTTATTTCGAACCCAGAGTCCCGCT**TCAGAAGA**ACTCGTCAAGAAGGC
GATAGAAGGCGATGCGCTGCGAATCGGGAGCGGCGATACCGTAAAGCACAGGAAGCGGT**CAGCCC**ATTTCG
CCGCCAAGCTCTTCAGCAATATCACGGGTAGCCAACGCTATGTCCTGATAGCGGTCCGCCACACCCAGCCG
GCCACAGTCGATGAATCCAGAAAAGCGGCCATTTTTCCACCATGATATTCGGCAAGCAGGCATCGCCATGGG
TCACGACGAGATCATCGCCGTCGGGCATGCGCGCCTTGAGCCTGGCGAACAGTTCGGCTGGCGCGAGCCCC
TGATGCTCTTCGTCCAGATCATCCTGATCGACAAGACCGGCTTCCATCCGAGTACGTGCTCGCTCGATGCG
ATGTTTCGCTTGGTGGT**CGAATGGGCAGGTAGCCGGATCAAGCGTATGCAGCCGCCGCATTGCATCAGCCA**
TGATGGATACTTTCTCGGCAGGAGCAAGGTGAGATGACAGGAGATCCTGCCCGGCACCTCGCCCAATAGC
AGCCAGTCCCTTCCCGCTTCAGTGACAACGTCGAGCACAGCTGCGCAAGGAACGCCCGTCTGGCCAGCCA
CGATAGCCCGCTGCCTCGTCTCGCAGTTCATT**CAGGGCACCGGACAGGT**CGGTCTTGACAAAAAGAACCG
GGCGCCCCTGCGCTGACAGCCGGAACACGGCGGCATCAGAGCAGCCGATTGTCTGTTGTGCCCAGTCATAG
CCGAATAGCCTCTCCACCCAAGCGGGCCGGAGAACCTGCGTGCAATCCATCTTGTTCAATGGCCGATCC**CAT**
GGTTTAGTTCCTCACCTTGTCGTATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAACACGTG
CTGCTGCAGGTCGAAAGGCCCGGAGATGAGGAAGAGGAGAACAGCGCGGCAGACGTGCGCTTTTGAAGCGT
GCAGAATGCCGGGCTCCGGAGGACCTTCGCGCCCCGCCCGCCCCCTGAGCCCCGCCCTGAGCCCCGCCCGG
ACCCACCCCTTCCCAGCCTCTGAGCCCAGAAAGCGAAGGAGCAAAGCTGCTATTGGCCGCTGCCCAAAGG
CCTACCCGCTTCCATTGCTCAGCGGTGCTGTCCATCTGCACGAGACTAGTGAGACGTGCTACTTCCATTTG
TCACGTCCTGCACGACGCGAGCTGCGGGGCGGGGGGAACTTCTGACTAGGGGAGGAGTAGAAGGTGGCG
CGAAGGGGCCACCAAAGAACGGAGCCGGTTGGCGCTACCGGTGGATGTGGAATGTGTGCGAGGCCAGAGGC
CACTTGTGTAGCGCCAAGTGCCAGCGGGGCTGCTAAAGCGCATGCTCCAGACTGCCTTGGGAAAAGCGCCT
CCCCTACCCGGTAGAAT**GAAGTTCCTATACTTTCTAGAGAATAGGAACTTC**GTTCGAAC**ATAACTTCGTAT**
AGCATAACATTATACGAAGTTATGGTACCTGCAGAATTCATGCATAAGCTTGGATCCGTTCTTCGGACGCCT
CGTCAACACCGTACG