```
5'  tggcgaatgggacgcgccctgtagcggcgcattaagcgcggcgggtgtggtggttacgcgcagcgtgacc
    ++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  70
3'  accgcttaccctgcgcgggacatcgccgcgtaattcgcgccgcccacaccaccaatgcgcgtcgcactgg

5'  gctacacttgccagcgccctagcgcccgctcctttcgctttcttcccttcctttctcgccacgttcgccg
    ++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  140
3'  cgatgtgaacggtcgcgggatcgcgggcgaggaaagcgaaagaagggaaggaaagagcggtgcaagcggc

5'  gctttccccgtcaagctctaaatcgggggctccctttagggttccgatttagtgctttacggcacctcga
    ++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  210
3'  cgaaaggggcagttcgagatttagcccccgagggaaatcccaaggctaaatcacgaaatgccgtggagct

                                          DraIII
                                          |
5'  ccccaaaaaacttgattagggtgatggttcacgtagtgggccatcgccctgatagacggtttttcgccct
    ++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  280
3'  ggggttttttgaactaatcccactaccaagtgcatcacccggtagcgggactatctgccaaaaagcggga

     AlwI'                                AlwI
     |                                    |
5'  ttgacgttggagtccacgttctttaatagtggactcttgttccaaactggaacaacactcaaccctatct
    ++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  350
3'  aactgcaacctcaggtgcaagaaattatcacctgagaacaaggtttgaccttgttgtgagttgggataga

5'  cggtctattcttttgatttataagggattttgccgatttcggcctattggttaaaaaatgagctgattta
    ++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  420
3'  gccagataagaaaactaaatattccctaaaacggctaaagccggataaccaatttttactcgactaaat

5'  acaaaaatttaacgcgaattttaacaaaatattaacgcttacaatttaggtggcacttttcggggaaatg
    ++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  490
3'  tgtttttaaattgcgcttaaaattgttttataattgcgaatgttaaatccaccgtgaaaagcccctttac

5'  tgcgcggaacccctatttgtttatttttctaaatacattcaaatatgtatccgctcatgagacaataacc
    ++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  560
3'  acgcgccttggggataaacaaataaaaagatttatgtaagtttatacataggcgagtactctgttattgg

5'  ctgataaatgcttcaataatattgaaaaaggaagagtatgagtattcaacatttccgtgtcgcccttatt
    ++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  630
3'  gactatttacgaagttattataactttttccttctcatactcataagttgtaaaggcacagcgggaataa
                                                 [Ap]

5'  ccctttttgcggcattttgccttcctgtttttgctcacccagaaacgctggtgaaagtaaaagatgctg
    ++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  700
3'  gggaaaaaacgccgtaaaacggaaggacaaaaacgagtgggtctttgcgaccactttcattttctacgac
    [Ap]

5'  aagatcagttgggtgcacgagtgggttacatcgaactggatctcaacagcggtaagatccttgagagttt
    ++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  770
3'  ttctagtcaacccacgtgctcacccaatgtagcttgacctagagttgtcgccattctaggaactctcaaa
    [Ap]
```

```
5'  tcgccccgaagaacgttttccaatgatgagcacttttaaagttctgctatgtggcgcggtattatcccgt
    ++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  840
3'  agcggggcttcttgcaaaaggttactactcgtgaaaatttcaagacgatacaccgcgccataatagggca
    ================================================================================
                                          Ap

                                                                      ScaI
5'  attgacgccgggcaagagcaactcggtcgccgcatacactattctcagaatgacttggttgagtactcac
    ++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  910
3'  taactgcggcccgttctcgttgagccagcggcgtatgtgataagagtcttactgaaccaactcatgagtg
    ================================================================================
                                          Ap

5'  cagtcacagaaaagcatcttacggatggcatgacagtaagagaattatgcagtgctgccataaccatgag
    ++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  980
3'  gtcagtgtcttttcgtagaatgcctaccgtactgtcattctcttaatacgtcacgacggtattggtactc
    ================================================================================
                                          Ap

                                                PvuI
5'  tgataacactgcggccaacttacttctgacaacgatcggaggaccgaaggagctaaccgctttttgcac
    ++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  1050
3'  actattgtgacgccggttgaatgaagactgttgctagcctcctggcttcctcgattggcgaaaaaacgtg
    ================================================================================
                                          Ap

5'  aacatgggggatcatgtaactcgccttgatcgttgggaaccggagctgaatgaagccataccaaacgacg
    ++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  1120
3'  ttgtacccccctagtacattgagcggaactagcaacccttggcctcgacttacttcggtatggtttgctgc
    ================================================================================
                                          Ap

                                    PstI
5'  agcgtgacaccacgatgcctgcagcaatggcaacaacgttgcgcaaactattaactggcgaactacttac
    ++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  1190
3'  tcgcactgtggtgctacggacgtcgttaccgttgttgcaacgcgtttgataattgaccgcttgatgaatg
    ================================================================================
                                          Ap

5'  tctagcttcccggcaacaattaatagactggatggaggcggataaagttgcaggaccacttctgcgctcg
    ++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  1260
3'  agatcgaagggccgttgttaattatctgacctacctccgcctatttcaacgtcctggtgaagacgcgagc
    ================================================================================
                                          Ap

                                                                      BsaI
5'  gcccttccggctggctggtttattgctgataaatctggagccggtgagcgtgggtctcgcggtatcattg
    ++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  1330
3'  cgggaaggccgaccgaccaaataacgactatttagacctcggccactcgcacccagagcgccatagtaac
    ================================================================================
                                          Ap
```

AhdI

5' cagcactggggccagatggtaagccctcccgtatcgtagttatctacacgacggggagtcaggcaactat 1400
3' gtcgtgaccccggtctaccattcgggagggcatagcatcaatagatgtgctgcccctcagtccgttgata

Ap

5' ggatgaacgaaatagacagatcgctgagataggtgcctcactgattaagcattggtaactgtcagaccaa 1470
3' cctacttgctttatctgtctagcgactctatccacggagtgactaattcgtaaccattgacagtctggtt

Ap

5' gtttactcatatatactttagattgatttaaaacttcatttttaatttaaaaggatctaggtgaagatcc 1540
3' caaatgagtatatatgaaatctaactaaattttgaagtaaaaattaaattttcctagatccacttctagg

5' tttttgataatctcatgaccaaaatcccttaacgtgagttttcgttccactgagcgtcagaccccgtaga 1610
3' aaaaactattagagtactggttttagggaattgcactcaaaagcaaggtgactcgcagtctggggcatct

5' aaagatcaaaggatcttcttgagatcctttttttctgcgcgtaatctgctgcttgcaaacaaaaaaacca 1680
3' tttctagtttcctagaagaactctaggaaaaaaagacgcgcattagacgacgaacgtttgtttttttggt

Ori

5' ccgctaccagcggtggtttgtttgccggatcaagagctaccaactctttttccgaaggtaactggcttca 1750
3' ggcgatggtcgccaccaaacaaacggcctagttctcgatggttgagaaaaaggcttccattgaccgaagt

Ori

5' gcagagcgcagataccaaatactgtccttctagtgtagccgtagttaggccaccacttcaagaactctgt 1820
3' cgtctcgcgtctatggtttatgacaggaagatcacatcggcatcaatccggtggtgaagttcttgagaca

Ori

5' agcaccgcctacatacctcgctctgctaatcctgttaccagtggctgctgccagtggcgataagtcgtgt 1890
3' tcgtggcggatgtatggagcgagacgattaggacaatggtcaccgacgacggtcaccgctattcagcaca

Ori

5' cttaccgggttggactcaagacgatagttaccggataaggcgcagcggtcgggctgaacggggggttcgt 1960
3' gaatggcccaacctgagttctgctatcaatggcctattccgcgtcgccagcccgacttgccccccaagca

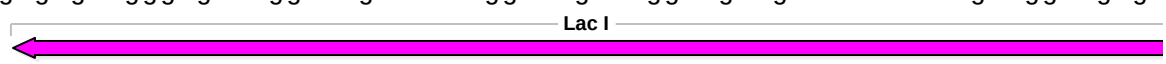
Ori

5' gcacacagcccagcttggagcgaacgacctacaccgaactgagatacctacagcgtgagctatgagaaag 2030
3' cgtgtgtcgggtcgaacctcgcttgctggatgtggcttgactctatggatgtcgcactcgatactctttc

Ori

```
5'  cgccacgcttcccgaagggagaaaggcggacaggtatccggtaagcggcagggtcggaacaggagagcgc
o   ++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  2100
3'  gcggtgcgaagggcttccctctttccgcctgtccataggccattcgccgtcccagccttgtcctctcgcg
o                                   Ori
o
5'  acgagggagcttccagggggaaacgcctggtatctttatagtcctgtcgggtttcgccacctctgacttg
o   ++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  2170
3'  tgctccctcgaaggtcccccttttgcggaccatagaaatatcaggacagcccaaagcggtggagactgaac
o                                   Ori
o
5'  agcgtcgatttttgtgatgctcgtcaggggggcggagcctatggaaaaacgcccagcaacgcggcctttt
o   ++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  2240
3'  tcgcagctaaaaacactacgagcagtcccccgcctcggatacctttttgcggtcgttgcgccggaaaaa
o                                   Ori
o
o                                       PciI
5'  acggttcctggccttttgctggccttttgctcacatgttctttcctgcgttatcccctgattctgtggat
o   ++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  2310
3'  tgccaaggaccggaaaacgaccggaaaacgagtgtacaagaaaggacgcaatagggactaagacaccta
o
5'  aaccgtattaccgcctttgagtgagctgataccgctcgccgcagccgaacgaccgagcgcagcgagtcag
o   ++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  2380
3'  ttggcataatggcggaaactcactcgactatggcgagcggcgtcggcttgctggctcgcgtcgctcagtc
o
o                 SapI
5'  tgagcgaggaagcggaagagcgcctgatgcggtattttctccttacgcatctgtgcggtatttcacaccg
o   ++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  2450
3'  actcgctccttcgccttctcgcggactacgccataaaagaggaatgcgtagacacgccataaagtgtggc
o
o                                                                   BstZ17I
5'  catatatggtgcactctcagtacaatctgctctgatgccgcatagttaagccagtatacactccgctatc
o   ++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  2520
3'  gtatataccacgtgagagtcatgttagacgagactacggcgtatcaattcggtcatatgtgaggcgatag
o
o                  Tth111I
5'  gctacgtgactgggtcatggctgcgccccgacacccgccaacacccgctgacgcgccctgacgggcttgt
o   ++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  2590
3'  cgatgcactgacccagtaccgacgcggggctgtgggcggttgtgggcgactgcgcgggactgcccgaaca
o
5'  ctgctcccggcatccgcttacagacaagctgtgaccgtctccgggagctgcatgtgtcagaggttttcac
o   ++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  2660
3'  gacgagggccgtaggcgaatgtctgttcgacactggcagaggccctcgacgtacacagtctccaaaagtg
o
5'  cgtcatcaccgaaacgcgcgaggcagctgcggtaaagctcatcagcgtggtcgtgaagcgattcacagat
o   ++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  2730
3'  gcagtagtggctttgcgcgctccgtcgacgccatttcgagtagtcgcaccagcacttcgctaagtgtcta
o
5'  gtctgcctgttcatccgcgtccagctcgttgagtttctccagaagcgttaatgtctggcttctgataaag
o   ++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  2800
3'  cagacggacaagtaggcgcaggtcgagcaactcaaagaggtcttcgcaattacagaccgaagactatttc
o
```

```
5'  cgggccatgttaagggcggtttttcctgtttggtcactgatgcctccgtgtaagggggatttctgttca
o   ++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  2870
3'  gcccggtacaattcccgccaaaaaaggacaaaccagtgactacggaggcacattccccctaaagacaagt
o

5'  tgggggtaatgataccgatgaaacgagagaggatgctcacgatacgggttactgatgatgaacatgcccg
o   ++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  2940
3'  acccccattactatggctactttgctctctcctacgagtgctatgcccaatgactactacttgtacgggc
o

5'  gttactggaacgttgtgagggtaaacaactggcggtatggatgcggcgggaccagagaaaaatcactcag
o   ++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  3010
3'  caatgaccttgcaacactcccatttgttgaccgccatacctacgccgccctggtctctttttagtgagtc
o

5'  ggtcaatgccagcgcttcgttaatacagatgtaggtgttccacagggtagccagcagcatcctgcgatgc
o   ++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  3080
3'  ccagttacggtcgcgaagcaattatgtctacatccacaaggtgtcccatcggtcgtcgtaggacgctacg
o

5'  agatccggaacataatggtgcagggcgctgacttccgcgtttccagactttacgaaacacggaaaccgaa
o   ++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  3150
3'  tctaggccttgtattaccacgtcccgcgactgaaggcgcaaaggtctgaaatgctttgtgcctttggctt
o

5'  gaccattcatgttgttgctcaggtcgcagacgttttgcagcagcagtcgcttcacgttcgctcgcgtatc
o   ++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  3220
3'  ctggtaagtacaacaacgagtccagcgtctgcaaaacgtcgtcgtcagcgaagtgcaagcgagcgcatag
o

5'  ggtgattcattctgctaaccagtaaggcaaccccgccagcctagccgggtcctcaacgacaggagcacga
o   ++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  3290
3'  ccactaagtaagacgattggtcattccgttggggcggtcggatcggcccaggagttgctgtcctcgtgct
o

o           FspAI
                |
5'  tcatgcgcacccgtggggccgccatgccggcgataatggcctgcttctcgccgaaacgtttggtggcggg
o   ++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  3360
3'  agtacgcgtgggcaccccggcggtacggccgctattaccggacgaagagcggctttgcaaaccaccgccc
o

5'  accagtgacgaaggcttgagcgagggcgtgcaagattccgaataccgcaagcgacaggccgatcatcgtc
o   ++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  3430
3'  tggtcactgcttccgaactcgctcccgcacgttctaaggcttatggcgttcgctgtccggctagtagcag
o

5'  gcgctccagcgaaagcggtcctcgccgaaaatgacccagagcgctgccggcacctgtcctacgagttgca
o   ++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  3500
3'  cgcgaggtcgctttcgccaggagcggcttttactgggtctcgcgacggccgtggacaggatgctcaacgt
o

5'  tgataaagaagacagtcataagtgcggcgacgatagtcatgccccgcgcccaccggaaggagctgactgg
o   ++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  3570
3'  actatttcttctgtcagtattcacgccgctgctatcagtacggggcgcgggtggccttcctcgactgacc
o

5'  gttgaaggctctcaagggcatcggtcgagatcccggtgcctaatgagtgagctaacttacattaattgcg
o   ++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  3640
3'  caacttccgagagttcccgtagccagctctagggccacggattactcactcgattgaatgtaattaacgc
o

5'  ttgcgctcactgcccgctttccagtcgggaaacctgtcgtgccagctgcattaatgaatcggccaacgcg
o   ++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  3710
3'  aacgcgagtgacgggcgaaaggtcagccctttggacagcacggtcgacgtaattacttagccggttgcgc
o
```



```
5'  cggggagaggcggtttgcgtattgggcgccagggtggtttttcttttcaccagtgagacgggcaacagct
    ++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  3780
3'  gcccctctccgccaaacgcataacccgcggtcccaccaaaaagaaaagtggtcactctgcccgttgtcga
                                   Lac I

5'  gattgcccttcaccgcctggccctgagagagttgcagcaagcggtccacgctggtttgccccagcaggcg
    ++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  3850
3'  ctaacgggaagtggcggaccgggactctctcaacgtcgttcgccaggtgcgaccaaacggggtcgtccgc
                                   Lac I

5'  aaaatcctgtttgatggtggttaacggcgggatataacatgagctgtcttcggtatcgtcgtatcccact
    ++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  3920
3'  ttttaggacaaactaccaccaattgccgccctatattgtactcgacagaagccatagcagcatagggtga
                                   Lac I

                                          BssHII
5'  accgagatgtccgcaccaacgcgcagcccggactcggtaatggcgcgcattgcgcccagcgccatctgat
    ++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  3990
3'  tggctctacaggcgtggttgcgcgtcgggcctgagccattaccgcgcgtaacgcgggtcgcggtagacta
                                   Lac I

5'  cgttggcaaccagcatcgcagtgggaacgatgccctcattcagcatttgcatggtttgttgaaaaccgga
    ++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  4060
3'  gcaaccgttggtcgtagcgtcacccttgctacgggagtaagtcgtaaacgtaccaaacaacttttggcct
                                   Lac I

5'  catggcactccagtcgccttcccgttccgctatcggctgaatttgattgcgagtgagatatttatgccag
    ++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  4130
3'  gtaccgtgaggtcagcggaagggcaaggcgatagccgacttaaactaacgctcactctataaatacggtc
                                   Lac I

                                          PspOMI
                                             ApaI                   BstEII
5'  ccagccagacgcagacgcgccgagacagaacttaatgggcccgctaacagcgcgatttgctggtgaccca
    ++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  4200
3'  ggtcggtctgcgtctgcgcggctctgtcttgaattacccgggcgattgtcgcgctaaacgaccactgggt
                                   Lac I

5'  atgcgaccagatgctccacgcccagtcgcgtaccgtcttcatgggagaaaataatactgttgatgggtgt
    ++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  4270
3'  tacgctggtctacgaggtgcgggtcagcgcatggcagaagtaccctcttttattatgacaactacccaca
                                   Lac I

5'  ctggtcagagacatcaagaaataacgccggaacattagtgcaggcagcttccacagcaatggcatcctgg
    ++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  4340
3'  gaccagtctctgtagttctttattgcggccttgtaatcacgtccgtcgaaggtgtcgttaccgtaggacc
                                   Lac I
```

BclI MluI

5' tcatccagcggatagttaatgatcagcccactgacgcgttgcgcgagaagattgtgcaccgccgctttac 4410
3' agtaggtcgcctatcaattactagtcgggtgactgcgcaacgcgctcttctaacacgtggcggcgaaatg

Lac I

5' aggcttcgacgccgcttcgttctaccatcgacaccaccacgctggcacccagttgatcggcgcgagattt 4480
3' tccgaagctgcggcgaagcaagatggtagctgtggtggtgcgaccgtgggtcaactagccgcgctctaaa

Lac I

5' aatcgccgcgacaatttgcgacggcgcgtgcagggccagactggaggtggcaacgccaatcagcaacgac 4550
3' ttagcggcgctgttaaacgctgccgcgcacgtcccggtctgacctccaccgttgcggttagtcgttgctg

Lac I

5' tgtttgcccgccagttgttgtgccacgcggttgggaatgtaattcagctccgccatcgccgcttccactt 4620
3' acaaacgggcggtcaacaacacggtgcgccaacccttacattaagtcgaggcggtagcggcgaaggtgaa

Lac I

5' tttcccgcgttttcgcagaaacgtggctggcctggttcaccacgcgggaaacggtctgataagagacacc 4690
3' aaagggcgcaaaagcgtctttgcaccgaccggaccaagtggtgcgccctttgccagactattctctgtgg

5' ggcatactctgcgacatcgtataacgttactggtttcacattcaccaccctgaattgactctcttccggg 4760
3' ccgtatgagacgctgtagcatattgcaatgaccaaagtgtaagtggtgggacttaactgagagaaggccc

5' cgctatcatgccataccgcgaaaggttttgcgccattcgatggtgtccgggatctcgacgctctccctta 4830
3' gcgatagtacggtatggcgctttccaaaacgcggtaagctaccacaggccctagagctgcgagagggaat

EcoNI

5' tgcgactcctgcattaggaagcagcccagtagtaggttgaggccgttgagcaccgccgccgcaaggaatg 4900
3' acgctgaggacgtaatccttcgtcgggtcatcatccaactccggcaactcgtggcggcggcgttccttac

SphI

5' gtgcatgcaaggagatggcgcccaacagtcccccggccacggggcctgccaccatacccacgccgaaaca 4970
3' cacgtacgttcctctaccgcgggttgtcagggggccggtgccccggacggtggtatgggtgcggctttgt

5' agcgctcatgagcccgaagtggcgagcccgatcttccccatcggtgatgtcggcgatataggcgccagca 5040
3' tcgcgagtactcgggcttcaccgctcgggctagaaggggtagccactacagccgctatatccgcggtcgt

SgrAI

5' accgcacctgtggcgccggtgatgccggccacgatgcgtccggcgtagaggatcgagatcgatctcgatc 5110
3' tggcgtggacaccgcggccactacggccggtgctacgcaggccgcatctcctagctctagctagagctag

XbaI

5' ccgcgaaattaatacgactcactatagggaattgtgagcggataacaattcccctctagaaataatttt 5180
3' ggcgctttaattatgctgagtgatatcccttaacactcgcctattgttaaggggagatctttattaaaa

T7 Promoter — Lac Operator

BarI — BarI'

5' gtttaactttaagaaggagatatacatatgagcgataaaattattcacctgactgacgacagttttgaca 5250
3' caaattgaaattcttcctctatatgtatactcgctattttaataagtggactgactgctgtcaaaactgt

Trx Tag

M S D K I I H L T D D S F D
1 2 3 4 5 6 7 8 9 10 11 12 13 14

RsrII

5' cggatgtactcaaagcggacggggcgatcctcgtcgatttctgggcagagtggtgcggtccgtgcaaaat 5320
3' gcctacatgagtttcgcctgccccgctaggagcagctaaagacccgtctcaccacgccaggcacgtttta

Trx Tag

T D V L K A D G A I L V D F W A E W C G P C K M
15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38

5' gatcgccccgattctggatgaaatcgctgacgaatatcagggcaaactgaccgttgcaaaactgaacatc 5390
3' ctagcggggctaagacctactttagcgactgcttatagtcccgtttgactggcaacgttttgacttgtag

Trx Tag

I A P I L D E I A D E Y Q G K L T V A K L N I
39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61

5' gatcaaaaccctggcactgcgccgaaatatggcatccgtggtatcccgactctgctgctgttcaaaaacg 5460
3' ctagttttgggaccgtgacgcggctttataccgtaggcaccatagggctgagacgacgacaagtttttgc

Trx Tag

D Q N P G T A P K Y G I R G I P T L L L F K N
62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84

5' gtgaagtggcggcaaccaaagtgggtgcactgtctaaaggtcagttgaaagagttcctcgacgctaacct 5530
3' cacttcaccgccgttggtttcacccacgtgacagatttccagtcaactttctcaaggagctgcgattgga

Trx Tag

G E V A A T K V G A L S K G Q L K E F L D A N L
85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108

MscI

5' ggccggttctggttctggccatatgcaccatcatcatcatcattcttctggtctggtgccacgcggttct 5600
3' ccggccaagaccaagaccggtatacgtggtagtagtagtagtaagaagaccagaccacggtgcgccaaga

T...g | His6-tag | Thrombin

A G S G S G H M H H H H H H S S G L V P R G S
109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131

BstBI Acc65I KpnI

5' ggtatgaaagaaaccgctgctgctaaattcgaacgccagcacatggacagcccagatctgggtaccgacg 5670
3' ccatactttctttggcgacgacgatttaagcttgcggtcgtgtacctgtcgggtctagacccatggctgc

En...e

G M K E T A A A K F E R Q H M D S P D L G T D
132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154

5' acgacgacaaggccatggcgatATCAACAAGTTTGTACAAAAAAGCAGGCTctgaaaacttgtactttca 5740
3' tgctgctgttccggtaccgctaTAGTTGTTCAAACATGTTTTTTCGTCCGAgacttttgaacatgaaagt

Enterokinase | attB1 | TEV site

BGH89_B01c-pDONR221

D D D K A M A I S T S L Y K K A G S E N L Y F Q
155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178

5' aggcGAGGTTTCTGATTCAGGAGATGGAAAGGTTTTAATAAAGGGGAATAATGGTGTTTCCTTAGCATCA 5810
3' tccgCTCCAAAGACTAAGTCCTCTACCTTTCCAAAATTATTTCCCCTTATTACCACAAAGGAATCGTAGT

T... | BGH89_B01c

BGH89_B01c-pDONR221

G E V S D S G D G K V L I K G N N G V S L A S
179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201

5' GGATTTAACTATTATTTAAAAAATTATTGTAATGTAAGTTACAATCCAATAATGGGGTCTAATTTAAAAA
5880
3' CCTAAATTGATAATAAATTTTTTAATAACATTACATTCAATGTTAGGTTATTACCCCAGATTAAATTTTT

BGH89_B01c

BGH89_B01c-pDONR221

G F N Y Y L K N Y C N V S Y N P I M G S N L K
202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224

5' TGCCTGAAACAATGCCTTCTGTTGGGGAGAGAGTAGTTATAGATACTCCATATGAGCATAGATATGCATT
5950
3' ACGGACTTTGTTACGGAAGACAACCCCTCTCTCATCAATATCTATGAGGTATACTCGTATCTATACGTAA

BGH89_B01c

BGH89_B01c-pDONR221

M P E T M P S V G E R V V I D T P Y E H R Y A L
225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248

5' AAACTTTTGTACCTATTCTTATACAATGTCTTTCTGGGATTGGGATCAATATGAAGAATTCCTAGATTGG
6020
3' TTTGAAAACATGGATAAGAATATGTTACAGAAAGACCCTAACCCTAGTTATACTTCTTAAGGATCTAACC

BGH89_B01c

BGH89_B01c-pDONR221

N F C T Y S Y T M S F W D W D Q Y E E F L D W
249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271

5' TGTGCTATGAATGGGGTTAACTTAGTTTTAGATATAATTGGACAAGAAGAGGTTTTAAGAAGAACCTTAA
6090
3' ACACGATACTTACCCCAATTGAATCAAAATCTATATTAACCTGTTCTTCTCCAAAATTCTTCTTGGAATT

BGH89_B01c

BGH89_B01c-pDONR221

C A M N G V N L V L D I I G Q E E V L R R T L
272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294

BspMI

5' ATGAATTTGGATATAGCGATGAAGAGGTTAAAGAATTTATTTCAGGACCTGCATATTTTGCTTGGTTCTA
6160
3' TACTTAAACCTATATCGCTACTTCTCCAATTTCTTAAATAAAGTCCTGGACGTATAAAACGAACCAAGAT

BGH89_B01c

BGH89_B01c-pDONR221

N E F G Y S D E E V K E F I S G P A Y F A W F Y
295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318

5' TATGCAAAATATGACTGGATTTGGAGGTCCACTTCCTAACGATTGGTTTGAGCAAAGAGCAGAGCTTGGA 6230
3' ATACGTTTTATACTGACCTAAACCTCCAGGTGAAGGATTGCTAACCAAACTCGTTTCTCGTCTCGAACCT

BGH89_B01c

BGH89_B01c-pDONR221

M Q N M T G F G G P L P N D W F E Q R A E L G
319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341

5' AGAAAAATGCATGATAGAATGCAAAGTTTTGGAATAAATCCTGTTCTTCAAGGATACTCAGGTATGGTTC 6300
3' TCTTTTTACGTACTATCTTACGTTTCAAAACCTTATTTAGGACAAGAAGTTCCTATGAGTCCATACCAAG

BGH89_B01c

BGH89_B01c-pDONR221

R K M H D R M Q S F G I N P V L Q G Y S G M V
342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364

5' CAAGAGACTTTAAGGAAAAAAATCAAGAAGCACAAACTATTTCTCAAGGGGGATGGTGTGGCTTTGATAG 6370
3' GTTCTCTGAAATTCCTTTTTTTAGTTCTTCGTGTTTGATAAAGAGTTCCCCCTACCACACCGAAACTATC

BGH89_B01c

BGH89_B01c-pDONR221

P R D F K E K N Q E A Q T I S Q G G W C G F D R
365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388

5' ACCAGATATGCTTAAAACATATGTAAATGAAGGGGAGGCAGATTATTTCCAAAAGGTTGCAGATGTTTTC 6440
3' TGGTCTATACGAATTTTGTATACATTTACTTCCCCTCCGTCTAATAAAGGTTTTCCAACGTCTACAAAAG

BGH89_B01c

BGH89_B01c-pDONR221

P D M L K T Y V N E G E A D Y F Q K V A D V F
389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411

5' TATGAAAAGCAAAAAGAAGTGTTTGGAGATGTAACTAACTTCTATGGAGTAGATCCATTCCATGAGGGAG 6510
3' ATACTTTTCGTTTTTCTTCACAAACCTCTACATTGATTGAAGATACCTCATCTAGGTAAGGTACTCCCTC

BGH89_B01c

BGH89_B01c-pDONR221

Y E K Q K E V F G D V T N F Y G V D P F H E G
412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434

5' GAAATACTGGTGATTTAGATAATGGAAAGATATATGAAATAATTCAAAATAAGATGATAGAGCATGATAA 6580
3' CTTTATGACCACTAAATCTATTACCTTTCTATATACTTTATTAAGTTTTATTCTACTATCTCGTACTATT

BGH89_B01c

BGH89_B01c-pDONR221

G N T G D L D N G K I Y E I I Q N K M I E H D N
435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458

5' TGATGCAGTTTGGGTTATACAAAACTGGCAAGGAAATCCTAGTAATAATAAGTTAGAAGGATTAACTAAA 6650
3' ACTACGTCAAACCCAATATGTTTTGACCGTTCCTTTAGGATCATTATTATTCAATCTTCCTAATTGATTT

BGH89_B01c

BGH89_B01c-pDONR221

D A V W V I Q N W Q G N P S N N K L E G L T K
459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481

5' AAAGACCAAGCTATGGTATTAGATCTTTTCTCAGAGGTAAGTCCAGACTGGAATAGATTAGAGGAAAGAG 6720
3' TTTCTGGTTCGATACCATAATCTAGAAAAGAGTCTCCATTCAGGTCTGACCTTATCTAATCTCCTTTCTC

BGH89_B01c

BGH89_B01c-pDONR221

K D Q A M V L D L F S E V S P D W N R L E E R
482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504

5' ATTTACCATGGATTTGGAATATGCTTCATAATTTTGGAGGAAGAATGGGAATGGATGCAGCACCAGAAAA 6790
3' TAAATGGTACCTAAACCTTATACGAAGTATTAAAACCTCCTTCTTACCCTTACCTACGTCGTGGTCTTTT

BGH89_B01c

BGH89_B01c-pDONR221

D L P W I W N M L H N F G G R M G M D A A P E K
505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528

5' ACTTGCAACAGAAATTCCTAAAGCCTTAGCTAATTCAGAGCATATGGTTGGTATAGGTATTACACCAGAG 6860
3' TGAACGTTGTCTTTAAGGATTTCGGAATCGATTAAGTCTCGTATACCAACCATATCCATAATGTGGTCTC

BGH89_B01c

BGH89_B01c-pDONR221

L A T E I P K A L A N S E H M V G I G I T P E
529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551

5' GCTATAAATACAAATCCTTTAGCGTATGAACTTTTATTTGACATGGCTTGGACTAGAGATCAAATAAACT 6930
3' CGATATTTATGTTTAGGAAATCGCATACTTGAAAATAAACTGTACCGAACCTGATCTCTAGTTTATTTGA

BGH89_B01c

BGH89_B01c-pDONR221

A I N T N P L A Y E L L F D M A W T R D Q I N
552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574

5' TTAGAACATGGACTGAAGATTATATTGAAAGAAGATATGGTAAGACTAATAAAGAGATTTTAGAGGCTTG 7000
3' AATCTTGTACCTGACTTCTAATATAACTTTCTTCTATACCATTCTGATTATTTCTCTAAAATCTCCGAAC

BGH89_B01c

BGH89_B01c-pDONR221

F R T W T E D Y I E R R Y G K T N K E I L E A W
575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598

5' GAATATTATATTAGACACTGCTTACAAAAAGAGAAATGATTATTACCAAGGTGCTGCTGAGTCTATTATA 7070
3' CTTATAATATAATCTGTGACGAATGTTTTTCTCTTTACTAATAATGGTTCCACGACGACTCAGATAATAT

BGH89_B01c

BGH89_B01c-pDONR221

N I I L D T A Y K K R N D Y Y Q G A A E S I I
599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621

5' AATGCAAGACCAGGCTTTGGAATAAAAAGTGCTTCAACTTGGGGACATTCAAAAATAGTTTATGATAAGA 7140
3' TTACGTTCTGGTCCGAAACCTTATTTTTCACGAAGTTGAACCCCTGTAAGTTTTTATCAAATACTATTCT

BGH89_B01c

BGH89_B01c-pDONR221

N A R P G F G I K S A S T W G H S K I V Y D K
622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644

5' GTGAATTTGAAAAAGCTATTGAAATATTTGCAAAGAACTATGATGAATTCAAAGATAGTGATGCATTCCT 7210
3' CACTTAAACTTTTTCGATAACTTTATAAACGTTTCTTGATACTACTTAAGTTTCTATCACTACGTAAGGA

BGH89_B01c

BGH89_B01c-pDONR221

S E F E K A I E I F A K N Y D E F K D S D A F L
645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668

5' TTATGACTTTGCTGATATATTAAAACAATTATTAGCTAACTCAGCTCAAGAATATTATGAGGTTATGTGT 7280
3' AATACTGAAACGACTATATAATTTTGTTAATAATCGATTGAGTCGAGTTCTTATAATACTCCAATACACA

BGH89_B01c

BGH89_B01c-pDONR221

Y D F A D I L K Q L L A N S A Q E Y Y E V M C
669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691

SwaI

5' AATGCTTATAATAATGGAAATGGAGAGAAATTTAAATTTGTTTCAGGAAAGTTTTTAGAACTTATAAAGC 7350
3' TTACGAATATTATTACCTTTACCTCTCTTTAAATTTAAACAAAGTCCTTTCAAAAATCTTGAATATTTCG

BGH89_B01c

BGH89_B01c-pDONR221

N A Y N N G N G E K F K F V S G K F L E L I K
692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714

5' TTCAAGAGAGAGTATTATCAACTAGACCTGAATTCTTAATAGGAAATTGGATTGAAGATGCTAGAACTAT 7420
3' AAGTTCTCTCTCATAATAGTTGATCTGGACTTAAGAATTATCCTTTAACCTAACTTCTACGATCTTGATA

BGH89_B01c

BGH89_B01c-pDONR221

L Q E R V L S T R P E F L I G N W I E D A R T M
715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738

AflII

5' GCTTAAGGACTCTGATGATTGGACTAAGGACTTATTTGAATTTAATGCTAGAGCTTTAGTAACAACTTGG 7490
3' CGAATTCCTGAGACTACTAACCTGATTCCTGAATAAACTTAAATTACGATCTCGAAATCATTGTTGAACC

BGH89_B01c

BGH89_B01c-pDONR221

L K D S D D W T K D L F E F N A R A L V T T W
739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761

5' GGTTCAAGAAATAATGCTGACGGTGGTGGGCTTAAAGATTATTCTAATAGACAATGGTCAGGATTAACTG 7560
3' CCAAGTTCTTTATTACGACTGCCACCACCCGAATTTCTAATAAGATTATCTGTTACCAGTCCTAATTGAC

BGH89_B01c

BGH89_B01c-pDONR221

G S R N N A D G G G L K D Y S N R Q W S G L T
762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784

5' AGGATTATTATTATGCTAGATGGGAAAAATGGATAAATGGATTACAAGCTGAATTAGATGGAGGAGCTAA 7630
3' TCCTAATAATAATACGATCTACCCTTTTTACCTATTTACCTAATGTTCGACTTAATCTACCTCCTCGATT

BGH89_B01c

BGH89_B01c-pDONR221

E D Y Y Y A R W E K W I N G L Q A E L D G G A K
785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808

BseRI

5' AGCCCCTAATATAGATTGGTTTAAGATGGAGTATGACTGGGTAAACAAGAAGAGTGATACAGATAAGCTT 7700
3' TCGGGGATTATATCTAACCAAATTCTACCTCATACTGACCCATTTGTTCTTCTCACTATGTCTATTCGAA

BGH89_B01c

BGH89_B01c-pDONR221

A P N I D W F K M E Y D W V N K K S D T D K L
809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831

5' TATCCTACAGAAGCTTCTAATGAAAACTTAGGAGAACTTGCTAAGATAGCCATGGAAAGTTATTCAGTAA 7770
3' ATAGGATGTCTTCGAAGATTACTTTTGAATCCTCTTGAACGATTCTATCGGTACCTTTCAATAAGTCATT

BGH89_B01c

BGH89_B01c-pDONR221

Y P T E A S N E N L G E L A K I A M E S Y S V
832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854

BamHI

5' CTAATATGGATAAAATATTAGGTGAAAATGAATAGGACCCAGCTTTCTTGTACAAAGTGGTTGATatcgg 7840
3' GATTATACCTATTTTATAATCCACTTTTACTTATCCTGGGTCGAAAGAACATGTTTCACCAACTAtagcc

BGH89_B01c

attB2

BGH89_B01c-pDONR221

T N M D K I L G E N E .
855 856 857 858 859 860 861 862 863 864 865 866

EcoICRI
SacI
SalI
NotI
EagI
PspXI
XhoI
AvaI
BmeT110I

5' atccgaattcgagctccgtcgacaagcttgcggccgcactcgagcaccaccaccaccaccactgagatcc 7910
3' taggcttaagctcgaggcagctgttcgaacgccggcgtgagctcgtggtggtggtggtggtgactctagg

BlpI

5' ggctgctaacaaagcccgaaaggaagctgagttggctgctgccaccgctgagcaataactagcataaccc 7980
3' ccgacgattgtttcgggctttccttcgactcaaccgacgacggtggcgactcgttattgatcgtattggg

T7 Terminator

5' cttggggcctctaaacgggtcttgaggggttttttgctgaaaggaggaactatatccggat

8041

3' gaaccccggagatttgcccagaactccccaaaaaacgactttcctccttgatataggccta

T7 Terminator