```
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'  ttgtaccccctagtacattgagcggaactagcaacccttggcctcgacttacttcggtatggtttgctgc
    ================================================================================
                                          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
    ++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  2100
3'  gcggtgcgaagggcttccctctttccgcctgtccataggccattcgccgtcccagccttgtcctctcgcg
                                   Ori

5'  acgagggagcttccagggggaaacgcctggtatctttatagtcctgtcgggtttcgccacctctgacttg
    ++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  2170
3'  tgctccctcgaaggtcccccttttgcggaccatagaaatatcaggacagcccaaagcggtggagactgaac
                                   Ori

5'  agcgtcgatttttgtgatgctcgtcaggggggcggagcctatggaaaaacgcccagcaacgcggcctttt
    ++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  2240
3'  tcgcagctaaaaacactacgagcagtcccccgcctcggatacctttttgcggtcgttgcgccggaaaaa
                                   Ori

                                          PciI
5'  acggttcctggccttttgctggccttttgctcacatgttctttcctgcgttatcccctgattctgtggat
    ++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  2310
3'  tgccaaggaccggaaaacgaccggaaaacgagtgtacaagaaaggacgcaatagggactaagacaccta

5'  aaccgtattaccgcctttgagtgagctgataccgctcgccgcagccgaacgaccgagcgcagcgagtcag
    ++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  2380
3'  ttggcataatggcggaaactcactcgactatggcgagcggcgtcggcttgctggctcgcgtcgctcagtc

                SapI
5'  tgagcgaggaagcggaagagcgcctgatgcggtattttctccttacgcatctgtgcggtatttcacaccg
    ++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  2450
3'  actcgctccttcgccttctcgcggactacgccataaaagaggaatgcgtagacacgccataaagtgtggc

5'  catatatggtgcactctcagtacaatctgctctgatgccgcatagttaagccagtatacactccgctatc
    ++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  2520
3'  gtatataccacgtgagagtcatgttagacgagactacggcgtatcaattcggtcatatgtgaggcgatag

                 Tth111I
5'  gctacgtgactgggtcatggctgcgccccgacacccgccaacacccgctgacgcgccctgacgggcttgt
    ++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  2590
3'  cgatgcactgacccagtaccgacgcggggctgtgggcggttgtgggcgactgcgcgggactgcccgaaca

5'  ctgctcccggcatccgcttacagacaagctgtgaccgtctccgggagctgcatgtgtcagaggttttcac
    ++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  2660
3'  gacgagggccgtaggcgaatgtctgttcgacactggcagaggccctcgacgtacacagtctccaaaagtg

5'  cgtcatcaccgaaacgcgcgaggcagctgcggtaaagctcatcagcgtggtcgtgaagcgattcacagat
    ++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  2730
3'  gcagtagtggctttgcgcgctccgtcgacgccatttcgagtagtcgcaccagcacttcgctaagtgtcta

5'  gtctgcctgttcatccgcgtccagctcgttgagtttctccagaagcgttaatgtctggcttctgataaag
    ++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  2800
3'  cagacggacaagtaggcgcaggtcgagcaactcaaagaggtcttcgcaattacagaccgaagactatttc
```

```
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

                         Bpu10I
o                        |
5'   gaccattcatgttgttgctcaggtcgcagacgttttgcagcagcagtcgcttcacgttcgctcgcgtatc
o    ++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  3220
3'   ctggtaagtacaacaacgagtccagcgtctgcaaaacgtcgtcgtcagcgaagtgcaagcgagcgcatag
o

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

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

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

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

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

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

```
5'  ttgcgctcactgcccgctttccagtcgggaaacctgtcgtgccagctgcattaatgaatcggccaacgcg
                                                                          3710
3'  aacgcgagtgacgggcgaaaggtcagccctttggacagcacggtcgacgtaattacttagccggttgcgc
                                  Lac I

5'  cggggagaggcggtttgcgtattgggcgccagggtggtttttcttttcaccagtgagacgggcaacagct
                                                                          3780
3'  gcccctctccgccaaacgcataacccgcggtcccaccaaaaagaaaagtggtcactctgcccgttgtcga
                                  Lac I

5'  gattgcccttcaccgcctggccctgagagagttgcagcaagcggtccacgctggtttgccccagcaggcg
                                                                          3850
3'  ctaacgggaagtggcggaccgggactctctcaacgtcgttcgccaggtgcgaccaaacggggtcgtccgc
                                  Lac I

                          HpaI
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

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
```

BglII 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
```

NcoI

```
5'  acgacgacaaggccatggcgatATCAACAAGTTTGTACAAAAAAGCAGGCTctgaaaacttgtactttca
                                                                          5740
3'  tgctgctgttccggtaccgctaTAGTTGTTCAAACATGTTTTTTCGTCCGAgacttttgaacatgaaagt
```

Enterokinase | attB1 | TEV site

BGT01_T01a-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'  aggcGCAAACGTACTCGTAATAAATTTCCCTGGAGAAGGTCATATAAATCCGACTTTAGCTATTGTAAGT
                                                                          5810
3'  tccgCGTTTGCATGAGCATTATTTAAAGGGACCTCTTCCAGTATATTTAGGCTGAAATCGATAACATTCA
```

T... BGT01_T01a

BGT01_T01a-pDONR221

```
G  A  N  V  L  V  I  N  F  P  G  E  G  H  I  N  P  T  L  A  I  V  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' GAGTTAATTCGGCGAGGGGAGACAGTTGTTTCGTATTGTATTGAAGATTATAGAAAGAAGATTGAAGCAA 5880
3' CTCAATTAAGCCGCTCCCCTCTGTCAACAAAGCATAACATAACTTCTAATATCTTTCTTCTAACTTCGTT

BGT01_T01a

BGT01_T01a-pDONR221

E L I R R G E T V V S Y C I E D Y R K K I E A
202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224

5' CAGGTGCACAATTCCGAGTGTTTGAGAATTTCCTCTCTCAAATTAATATTATGGAGCGAGTAAATGAAGG 5950
3' GTCCACGTGTTAAGGCTCACAAACTCTTAAAGGAGAGAGTTTAATTATAATACCTCGCTCATTTACTTCC

BGT01_T01a

BGT01_T01a-pDONR221

T G A Q F R V F E N F L S Q I N I M E R V N E G
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' TGGGAGTCCTTTGACGATGCTGTCTCATATGATGGAAGCATCAGAACGTATTGTTACTCAAATTGTAGAA 6020
3' ACCCTCAGGAAACTGCTACGACAGAGTATACTACCTTCGTAGTCTTGCATAACAATGAGTTTAACATCTT

BGT01_T01a

BGT01_T01a-pDONR221

G S P L T M L S H M M E A S E R I V T Q I V E
249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271

5' GAAACAAAAGGGGAAAAGTACGATTATTTGATATATGATAATCACTTTCCAGTAGGACGTATTATAGCCA 6090
3' CTTTGTTTTCCCCTTTTCATGCTAATAAACTATATACTATTAGTGAAAGGTCATCCTGCATAATATCGGT

BGT01_T01a

BGT01_T01a-pDONR221

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

5' ATGTTTTAAAGTTACCTAGTGTTTCTTCTTGTACAACGTTTGCTTTTAATCAGTACATTACTTTTAACGA 6160
3' TACAAAATTTCAATGGATCACAAAGAAGAACATGTTGCAAACGAAAATTAGTCATGTAATGAAAATTGCT

BGT01_T01a

BGT01_T01a-pDONR221

N V L K L P S V S S C T T F A F N Q Y I T F N D
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' TGAACATGAATCAAGAGAAGTAGATGAAACGAATCCATTGTATCAATCTTGTTTAGCGGGAATGGAAAAA 6230
3' ACTTGTACTTAGTTCTCTTCATCTACTTTGCTTAGGTAACATAGTTAGAACAAATCGCCCTTACCTTTTT

BGT01_T01a

BGT01_T01a-pDONR221

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

5' TGGAACAAACAGTATGGAATGAAATGTAATAGTATGTATGATATTATGAACCATCCTGGTGATATTACAA 6300
3' ACCTTGTTTGTCATACCTTACTTTACATTATCATACATACTATAATACTTGGTAGGACCACTATAATGTT

BGT01_T01a

BGT01_T01a-pDONR221

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

5' TTGTGTATACTTCAAAGGAATATCAGCCGCGTTCAGATGTATTCGATGAATCGTATAAGTTTGTTGGCCC 6370
3' AACACATATGAAGTTTCCTTATAGTCGGCGCAAGTCTACATAAGCTACTTAGCATATTCAAACAACCGGG

BGT01_T01a

BGT01_T01a-pDONR221

I V Y T S K E Y Q P R S D V F D E S Y K F V G P
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' ATCAATTGCTACTCGAAAAGAAGTAGGTAGCTTTCCTATGGAAGATTTAAAAGATGAAAAATTGATTTTC 6440
3' TAGTTAACGATGAGCTTTTCTTCATCCATCGAAAGGATACCTTCTAAATTTTCTACTTTTTAACTAAAAG

BGT01_T01a

BGT01_T01a-pDONR221

S I A T R K E V G S F P M E D L K D E K L I 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' ATTTCTATGGGAACAGTTTTTAATGAACAACCTGAGTTATATGAAAAATGTTTTGAAGCGTTTAAAGATG 6510
3' TAAAGATACCCTTGTCAAAAATTACTTGTTGGACTCAATATACTTTTTACAAAACTTCGCAAATTTCTAC

BGT01_T01a

BGT01_T01a-pDONR221

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

```
5'  TAGAAGCGACAGTCGTATTAGTTGTTGGTAAGAAGATAAATATAAGTCAATTTGAAAACATTCCGAATAA
                                                                          6580
3'  ATCTTCGCTGTCAGCATAATCAACAACCATTCTTCTATTTATATTCAGTTAAACTTTTGTAAGGCTTATT
```

BGT01_T01a

BGT01_T01a-pDONR221

```
V  E  A  T  V  V  L  V  V  G  K  K  I  N  I  S  Q  F  E  N  I  P  N  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
```

PsrI'

```
5'  CTTTAAGTTGTATAATTATGTGCCGCAATTAGAACTATTACAGTATGCTGATGTATTCGTAACACACGGC
                                                                          6650
3'  GAAATTCAACATATTAATACACGGCGTTAATCTTGATAATGTCATACGACTACATAAGCATTGTGTGCCG
```

BGT01_T01a

BGT01_T01a-pDONR221

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

PsrI

```
5'  GGTATGAATAGTTCAAGTGAAGCACTATATTACGGTGTCCCGTTAGTTGTAATTCCGGTAACAGGAGATC
                                                                          6720
3'  CCATACTTATCAAGTTCACTTCGTGATATAATGCCACAGGGCAATCAACATTAAGGCCATTGTCCTCTAG
```

BGT01_T01a

BGT01_T01a-pDONR221

```
G  M  N  S  S  S  E  A  L  Y  Y  G  V  P  L  V  V  I  P  V  T  G  D
482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504
```

```
5'  AGCCTTTAGTTGCGAAACGAGTAAATGAAGTAGGGGCTGGAATAAGGCTTAATCGCAAAGAATTAACTTC
                                                                          6790
3'  TCGGAAATCAACGCTTTGCTCATTTACTTCATCCCCGACCTTATTCCGAATTAGCGTTTCTTAATTGAAG
```

BGT01_T01a

BGT01_T01a-pDONR221

```
Q  P  L  V  A  K  R  V  N  E  V  G  A  G  I  R  L  N  R  K  E  L  T  S
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'  TGAAATGTTACGTGAGTCTGTAAAGAAAGTGATGGATGATGTAACGTTTAAGGAAAAAAGTCGTAAAGTT
                                                                          6860
3'  ACTTTACAATGCACTCAGACATTTCTTTCACTACCTACTACATTGCAAATTCCTTTTTTCAGCATTTCAA
```

BGT01_T01a

BGT01_T01a-pDONR221

```
E  M  L  R  E  S  V  K  K  V  M  D  D  V  T  F  K  E  K  S  R  K  V
529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551
```

5' GGAGAGTCACTTCGAAATGCTGGTGGTTATAATAGGGCAGTTGATGAAATATTAAAAATGAATTCATACT
6930
3' CCTCTCAGTGAAGCTTTACGACCACCAATATTATCCCGTCAACTACTTTATAATTTTTACTTAAGTATGA

BGT01_T01a

BGT01_T01a-pDONR221

G E S L R N A G G Y N R A V D E I L K M N S Y
552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574

EcoICRI
SacI
BamHI
SalI

5' CAAAACTTAAATAGGACCCAGCTTTCTTGTACAAAGTGGTTGATatcggatccgaattcgagctccgtcg
7000
3' GTTTTGAATTTATCCTGGGTCGAAAGAACATGTTTCACCAACTAtagcctaggcttaagctcgaggcagc

BGT01_T01a
attB2
BGT01_T01a-pDONR221

S K L K .
575 576 577 578 579

PspXI
XhoI
AvaI
NotI
EagI
HindIII
BmeT110I

5' acaagcttgcggccgcactcgagcaccaccaccaccaccactgagatccggctgctaacaaagcccgaaa
7070
3' tgttcgaacgccggcgtgagctcgtggtggtggtggtggtgactctaggccgacgattgtttcgggcttt

BlpI

5' ggaagctgagttggctgctgccaccgctgagcaataactagcataaccccttggggcctctaaacgggtc
7140
3' ccttcgactcaaccgacgacggtggcgactcgttattgatcgtattggggaaccccggagatttgcccag

T7 Terminator

5' ttgaggggttttttgctgaaaggaggaactatatccggat
7180
3' aactccccaaaaaacgactttcctccttgatataggccta

T7 Terminator