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

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

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

5'  ccccaaaaaacttgattagggtgatggttcacgtagtgggccatcgccctgatagacggtttttcgccct
    ++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  280
3'  ggggtttttgaactaatcccactaccaagtgcatcacccggtagcgggactatctgccaaaaagcggga

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

                          PsiI
                          |
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
o   ++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  1470
3'  cctacttgctttatctgtctagcgactctatccacggagtgactaattcgtaaccattgacagtctggtt
o   ==========================================================Ap==========

o
5'  gtttactcatatatactttagattgatttaaaacttcatttttaatttaaaaggatctaggtgaagatcc
o   ++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  1540
3'  caaatgagtatatatgaaatctaactaaattttgaagtaaaaattaaattttcctagatccacttctagg
o

5'  tttttgataatctcatgaccaaaatcccttaacgtgagttttcgttccactgagcgtcagaccccgtaga
o   ++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  1610
3'  aaaaactattagagtactggttttagggaattgcactcaaaagcaaggtgactcgcagtctggggcatct
o

5'  aaagatcaaaggatcttcttgagatccttttttctgcgcgtaatctgctgcttgcaaacaaaaaaacca
o   ++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  1680
3'  tttctagtttcctagaagaactctaggaaaaaaagacgcgcattagacgacgaacgtttgtttttttggt
o   ==========================Ori==========================================

o
5'  ccgctaccagcggtggtttgtttgccggatcaagagctaccaactctttttccgaaggtaactggcttca
o   ++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  1750
3'  ggcgatggtcgccaccaaacaaacggcctagttctcgatggttgagaaaaaggcttccattgaccgaagt
o   ==========================Ori==========================================

o
5'  gcagagcgcagataccaaatactgtccttctagtgtagccgtagttaggccaccacttcaagaactctgt
o   ++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  1820
3'  cgtctcgcgtctatggtttatgacaggaagatcacatcggcatcaatccggtggtgaagttcttgagaca
o   ==========================Ori==========================================

o
5'  agcaccgcctacatacctcgctctgctaatcctgttaccagtggctgctgccagtggcgataagtcgtgt
o   ++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  1890
3'  tcgtggcggatgtatggagcgagacgattaggacaatggtcaccgacgacggtcaccgctattcagcaca
o   ==========================Ori==========================================

o
5'  cttaccgggttggactcaagacgatagttaccggataaggcgcagcggtcgggctgaacggggggttcgt
o   ++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  1960
3'  gaatggcccaacctgagttctgctatcaatggcctattccgcgtcgccagcccgacttgccccccaagca
o   ==========================Ori==========================================

o
5'  gcacacagcccagcttggagcgaacgacctacaccgaactgagatacctacagcgtgagctatgagaaag
o   ++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  2030
3'  cgtgtgtcgggtcgaacctcgcttgctggatgtggcttgactctatggatgtcgcactcgatactctttc
o   ==========================Ori==========================================

o
5'  cgccacgcttcccgaagggagaaaggcggacaggtatccggtaagcggcagggtcggaacaggagagcgc
o   ++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  2100
3'  gcggtgcgaagggcttccctctttccgcctgtccataggccattcgccgtcccagccttgtcctctcgcg
o   ==========================Ori==========================================

o
```

```
5'  acgagggagcttccagggggaaacgcctggtatctttatagtcctgtcgggtttcgccacctctgacttg
o   ++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  2170
3'  tgctccctcgaaggtcccccttttgcggaccatagaaatatcaggacagcccaaagcggtggagactgaac
o                                   Ori
o
5'  agcgtcgatttttgtgatgctcgtcaggggggcggagcctatggaaaaacgccagcaacgcggcctttttt
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
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

          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
o   ++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  3710
3'  aacgcgagtgacgggcgaaaggtcagccctttggacagcacggtcgacgtaattacttagccggttgcgc
o           |                                   Lac I                      |
o           <==========================================================>
o
```

```
5'  cggggagaggcggtttgcgtattgggcgccagggtggtttttcttttcaccagtgagacgggcaacagct
o   ++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  3780
3'  gcccctctccgccaaacgcataacccgcggtcccaccaaaaagaaaagtggtcactctgcccgttgtcga
o                                   Lac I
o
5'  gattgcccttcaccgcctggccctgagagagttgcagcaagcggtccacgctggtttgccccagcaggcg
o   ++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  3850
3'  ctaacgggaagtggcggaccgggactctctcaacgtcgttcgccaggtgcgaccaaacggggtcgtccgc
o                                   Lac I
o
o                          HpaI
5'  aaaatcctgtttgatggtggttaacggcgggatataacatgagctgtcttcggtatcgtcgtatcccact
o   ++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  3920
3'  ttttaggacaaactaccaccaattgccgccctatattgtactcgacagaagccatagcagcatagggtga
o                                   Lac I
o
5'  accgagatgtccgcaccaacgcgcagcccggactcggtaatggcgcgcattgcgcccagcgccatctgat
o   ++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  3990
3'  tggctctacaggcgtggttgcgcgtcgggcctgagccattaccgcgcgtaacgcgggtcgcggtagacta
o                                   Lac I
o
5'  cgttggcaaccagcatcgcagtgggaacgatgccctcattcagcatttgcatggtttgttgaaaaccgga
o   ++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  4060
3'  gcaaccgttggtcgtagcgtcacccttgctacgggagtaagtcgtaaacgtaccaaacaacttttggcct
o                                   Lac I
o
5'  catggcactccagtcgccttcccgttccgctatcggctgaatttgattgcgagtgagatatttatgccag
o   ++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  4130
3'  gtaccgtgaggtcagcggaagggcaaggcgatagccgacttaaactaacgctcactctataaatacggtc
o                                   Lac I
o
o                                   PspOMI
                                      ApaI
5'  ccagccagacgcagacgcgccgagacagaacttaatgggcccgctaacagcgcgatttgctggtgaccca
o   ++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  4200
3'  ggtcggtctgcgtctgcgcggctctgtcttgaattacccgggcgattgtcgcgctaaacgaccactgggt
o                                   Lac I
o
5'  atgcgaccagatgctccacgcccagtcgcgtaccgtcttcatgggagaaaataatactgttgatgggtgt
o   ++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  4270
3'  tacgctggtctacgaggtgcgggtcagcgcatggcagaagtaccctcttttattatgacaactacccaca
o                                   Lac I
o
5'  ctggtcagagacatcaagaaataacgccggaacattagtgcaggcagcttccacagcaatggcatcctgg
o   ++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  4340
3'  gaccagtctctgtagttctttattgcggccttgtaatcacgtccgtcgaaggtgtcgttaccgtaggacc
o                                   Lac I
o
```

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

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 | 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

5' acgacgacaaggccatggcgatATCAACAAGTTTGTACAAAAAAGCAGGCTctgaaaacttgtactttca 5740
3' tgctgctgttccggtaccgctaTAGTTGTTCAAACATGTTTTTTCGTCCGAgacttttgaacatgaaagt

Enterokinase | attB1 | TEV site

BGH13_B04a-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' aggcACGACAACCCCCTGGTGGCGCGGCGCGGTTATTTATCAGGTATACCCCCGCAGTTTGATGGATACC 5810
3' tccgTGCTGTTGGGGGACCACCGCGCCGCGCCAATAAATAGTCCATATGGGGGCGTCAAACTACCTATGG

T... | BGH13_B04a

BGH13_B04a-pDONR221

G T T T P W W R G A V I Y Q V Y P R S L M D T
179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201

5' AATGGCGATGGCATTGGTGATATTCCCGGCATTATCAAAAAGCTCGACTACATCGCCAGCCTGGGTGTGG 5880
3' TTACCGCTACCGTAACCACTATAAGGGCCGTAATAGTTTTTCGAGCTGATGTAGCGGTCGGACCCACACC

BGH13_B04a

BGH13_B04a-pDONR221

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

5' ACGCCATTTGGGTATCGCCATTTTTCAAGTCGCCGATGAAAGACTTCGGTTACGACATTAGTGACTATCG 5950
3' TGCGGTAAACCCATAGCGGTAAAAAGTTCAGCGGCTACTTTCTGAAGCCAATGCTGTAATCACTGATAGC

BGH13_B04a

BGH13_B04a-pDONR221

D A I W V S P F F K S P M K D F G Y D I S D Y R
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' GGATATTGATCCCATCTTTGGAACCCTGAGTGATTTTGATGAGCTGATCAGCAAAGCCCATGCGCGCAAT 6020
3' CCTATAACTAGGGTAGAAACCTTGGGACTCACTAAAACTACTCGACTAGTCGTTTCGGGTACGCGCGTTA

BGH13_B04a

BGH13_B04a-pDONR221

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

5' ATCAAGATTATTATCGATCAGGTGCTCAGCCATACCTCGGACCAGCATGCCTGGTTTTTGGAGAGCCGCG 6090
3' TAGTTCTAATAATAGCTAGTCCACGAGTCGGTATGGAGCCTGGTCGTACGGACCAAAAACCTCTCGGCGC

BGH13_B04a

BGH13_B04a-pDONR221

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

5' AGTCGCGCGACAACCCGAAAGCTGATTGGTATGTGTGGGTAGATCCTCAACCCGACGGAACCCCCCCCAA 6160
3' TCAGCGCGCTGTTGGGCTTTCGACTAACCATACACACCCATCTAGGAGTTGGGCTGCCTTGGGGGGGGTT

BGH13_B04a

BGH13_B04a-pDONR221

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

MfeI

5' CAATTGGTTGGCAATTTTTGGGGGCTCTGCCTGGGAGTGGGAGCCGCGTCGTTGTCAATATTACCTGCAT 6230
3' GTTAACCAACCGTTAAAAACCCCCGAGACGGACCCTCACCCTCGGCGCAGCAACAGTTATAATGGACGTA

BGH13_B04a

BGH13_B04a-pDONR221

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

BspMI SwaI

5' AACTTTTTGCGTACCCAGCCAGATTTAAATTACCACTGTCCGCAAGTGCGCGAGCAAATTTTGCAGGAAG 6300
3' TTGAAAAACGCATGGGTCGGTCTAAATTTAATGGTGACAGGCGTTCACGCGCTCGTTTAAAACGTCCTTC

BGH13_B04a

BGH13_B04a-pDONR221

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

5' TGGAGTTCTGGTTGAAACGCGGCGTGGATGGTTTGCGCCTTGATGCCATTAACTTCTGCTTTCACGATAA 6370
3' ACCTCAAGACCAACTTTGCGCCGCACCTACCAAACGCGGAACTACGGTAATTGAAGACGAAAGTGCTATT

BGH13_B04a

BGH13_B04a-pDONR221

V E F W L K R G V D G L R L D A I N F C F H D K
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' GGAATTGCGTTCCAATCCTGCCAAGCCGGAAGCAGAGCGTAAGGGGCGAGGCTTTAAGGAAGACAACCCC 6440
3' CCTTAACGCAAGGTTAGGACGGTTCGGCCTTCGTCTCGCATTCCCCGCTCCGAAATTCCTTCTGTTGGGG

BGH13_B04a

BGH13_B04a-pDONR221

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

BsmI

5' TACGCATTCCAGAAACACATATTTGATAATACCCGCCCGGAAAACCTGGCTTTCCTGGAGTCCTTGCGCC 6510
3' ATGCGTAAGGTCTTTGTGTATAAACTATTATGGGCGGGCCTTTTGGACCGAAAGGACCTCAGGAACGCGG

BGH13_B04a

BGH13_B04a-pDONR221

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

XmaI
SmaI

5' AGTTAATGGATCGTTATCCCGGGACGGTAACCCTCGGTGAAATTTCAGCCGAGGACTCCCTGAAAACCAT 6580
3' TCAATTACCTAGCAATAGGGCCCTGCCATTGGGAGCCACTTTAAAGTCGGCTCCTGAGGGACTTTTGGTA

BGH13_B04a

BGH13_B04a-pDONR221

Q L M D R Y P G T V T L G E I S A E D S L K T M
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' GGCGGAATACACCTCCGGTAACGACCGTTTGCATATGGCGTATAGCTTTGAACTGCTGGTCGATAAGTTA 6650
3' CCGCCTTATGTGGAGGCCATTGCTGGCAAACGTATACCGCATATCGAAACTTGACGACCAGCTATTCAAT

BGH13_B04a

BGH13_B04a-pDONR221

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

BbvCI

5' TCCGGCAACTATATCCGCAATACGGTGGAAACCCTGGAAGCCAATTTGGCTGAGGGCTGGCCCTGTTGGT 6720
3' AGGCCGTTGATATAGGCGTTATGCCACCTTTGGGACCTTCGGTTAAACCGACTCCCGACCGGGACAACCA

BGH13_B04a

BGH13_B04a-pDONR221

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

5' CTATAGGCAACCACGATGTGGTGCGATTTATGTCGCGTTGGGGCGGTGAGGCACAATCCCCCGAACTCGC 6790
3' GATATCCGTTGGTGCTACACCACGCTAAATACAGCGCAACCCCGCCACTCCGTGTTAGGGGGCTTGAGCG

BGH13_B04a

BGH13_B04a-pDONR221

S I G N H D V V R F M S R W G G E A Q S P E L A
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' CAAAACCCTCAATGCCATGCTGCTTTCATTGCGCGGCAGTGTGTGTAGCTATCAGGGGGAAGAACTGGGT 6860
3' GTTTTGGGAGTTACGGTACGACGAAAGTAACGCGCCGTCACACACATCGATAGTCCCCCTTCTTGACCCA

BGH13_B04a

BGH13_B04a-pDONR221

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

5' TTGACAGAAGCGGATATCCAGCAGCATGAGTTGCAAGATCCCTATGGTATTACCTTCTGGCCTCGCTTCA 6930
3' AACTGTCTTCGCCTATAGGTCGTCGTACTCAACGTTCTAGGGATACCATAATGGAAGACCGGAGCGAAGT

BGH13_B04a

BGH13_B04a-pDONR221

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

PasI

5' AGGGGCGCGATGGATGTAGAACGCCAATGCCCTGGGACTCATCCCTGGAATACGCGGGATTTTCCAGTGC 7000
3' TCCCCGCGCTACCTACATCTTGCGGTTACGGGACCCTGAGTAGGGACCTTATGCGCCCTAAAAGGTCACG

BGH13_B04a

BGH13_B04a-pDONR221

K G R D G C R T P M P W D S S L E Y A G F S S A
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' AAAACCATGGCTACCCGTATCGGCAAAACACCTGGCGCTCTCCATCGATAAACAGGAAAGCAATCAGGCA 7070
3' TTTTGGTACCGATGGGCATAGCCGTTTTGTGGACCGCGAGAGGTAGCTATTTGTCCTTTCGTTAGTCCGT

BGH13_B04a

BGH13_B04a-pDONR221

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

5' TCTGTGCTGAACGCCTATCGCCATTTTATGCAGTGGCGTAAAACGCAAGCTGCCCTGCGCTTTGGTGGCA 7140
3' AGACACGACTTGCGGATAGCGGTAAAATACGTCACCGCATTTTGCGTTCGACGGGACGCGAAACCACCGT

BGH13_B04a

BGH13_B04a-pDONR221

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

5' TTGAATTTCTCTCCAGTAGTGAATCTCATTTGGTGTTTGTTCGTCGCCATGGCAATGAGGCGCTATTGTG 7210
3' AACTTAAAGAGAGGTCATCACTTAGAGTAAACCACAAACAAGCAGCGGTACCGTTACTCCGCGATAACAC

BGH13_B04a

BGH13_B04a-pDONR221

I E F L S S S E S H L V F V R R H G N E A L L C
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' TGGATTTAATTTCAGTGCCCAGCCTGCCGTTATTGAGTTGGGACAAGTGTTGGGCCTGGAGTCAATAACC 7280
3' ACCTAAATTAAAGTCACGGGTCGGACGGCAATAACTCAACCCTGTTCACAACCCGGACCTCAGTTATTGG

BGH13_B04a

BGH13_B04a-pDONR221

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

5' GGCCATGGTTTTGCAACGCAATCCTCCCGGTCAAACAATATTGGTTTGCCGGGGTATAGCGCGTTTTTTG 7350
3' CCGGTACCAAAACGTTGCGTTAGGAGGGCCAGTTTGTTATAACCAAACGGCCCCATATCGCGCAAAAAAC

BGH13_B04a

BGH13_B04a-pDONR221

G H G F A T Q S S R S N N I G L P G Y S A F F
692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714

EcoICRI SacI BamHI EcoRI SalI

5' CCAGCCTTGTCTAGGACCCAGCTTTCTTGTACAAAGTGGTTGATatcggatccgaattcgagctccgtcg 7420
3' GGTCGGAACAGATCCTGGGTCGAAAGAACATGTTTCACCAACTAtagcctaggcttaagctcgaggcagc

BGH13_B04a attB2

BGH13_B04a-pDONR221

A S L V .
715 716 717 718 719

HindIII NotI EagI PspXI XhoI

5' acaagcttgcggccgcactcgagcaccaccaccaccaccactgagatccggctgctaacaaagcccgaaa 7490
3' tgttcgaacgccggcgtgagctcgtggtggtggtggtggtgactctaggccgacgattgtttcgggcttt

5' ggaagctgagttggctgctgccaccgctgagcaataactagcataaccccttggggcctctaaacgggtc 7560
3' ccttcgactcaaccgacgacggtggcgactcgttattgatcgtattggggaaccccggagatttgcccag

T7 Terminator

5' ttgaggggttttttgctgaaaggaggaactatatccggat 7600
3' aactccccaaaaaacgactttcctccttgatataggccta

T7 Terminator