

	AscI	NdeI	BamHI	NotI	KpnI	SacI	ScaI			
	~~~~~	~~~~~	~~~~~	~~~~~	~~~~~	~~~~~	~~~~~			
1	CGCGCCCAT	ATGTCGACGG	ATCCCCGGGG	CGGCCGCGGG	TACCGAGCTC	TAAAAGGAAT	GTCTCCCTTG	CCAGTACTGC	TAGGGTTTTT	CTTCAAACCT
	GCGCGGGGTA	TACAGCTGCC	TAGGGGCCCC	GCCGGCGCCC	ATGGCTCGAG	ATTTTCCTTA	CAGAGGGAAC	GGTCATGACG	ATCCCAAAAA	GAAAGTTTGA
										ScaI
										~~~
101	ATGGAAGCCC	ATTCAAGCTG	CATATTACGA	TTTTGTTTTT	CGCTTTTAGA	AAGTGGTTTA	GATGAGATAA	TAGAAAAATT	CTTGATCTCC	GACAACGAGT
	TACCTTCGGG	TAAGTTCGAC	GTATAATGCT	AAAACAAAAA	GCGAAAATCT	TTCACCAAAT	CTACTCTATT	ATCTTTTTAA	GAAGTAGAGG	CTGTTGCTCA
										ScaI
										~~~
201	ACTTTTATTT	TTTTTGCTAA	TCACCTTACT	CAATATTAGC	TCGAAATCGT	AGAAACGTAG	ACGGGTGCGG	GATACCGAGT	GGTGTAGTTA	AGAATTTTTA
	TGAAAATAAA	AAAAACGATT	AGTGAAATGA	GTTATAATCG	AGCTTTAGCA	TCTTTGCATC	TGCCACGCC	CTATGGCTCA	CCACATCAAT	TCTTAAAAAT
301	TAAACCACGT	GGCCCAAAAA	TATGAACCCA	AAACGTTTAT	ACATGAGTAT	ACTTTAAGAA	GGCTATACCC	CTTCGTGTTA	GATGTAGTTT	TAGCTACCCA
	ATTTGGTGCA	CCGGGTTTTT	ATACTTGGGT	TTTGCAAATA	TGTACTCATA	TGAAATCTTT	CCGATATGGG	GAAGCACAAT	CTACATCAAA	ATCGATGGGT
401	ACCCGAGTCT	ATGAGCTTGA	CTTCAGATGT	AGAAGGCATF	AAATCGTTTT	GAATATTATA	TAAAAACGA	TGAAAATTAA	ATATTTAAAA	GCAATCATAAC
	TGGGCTCAGA	TACTCGAACT	GAAGTCTACA	TCTTCCGTAA	TTTAGCAAAA	CTTATAATAT	ATTTTTTGCT	ACTTTTAATT	TATAAATTTT	CGTTAGTATG
501	GCTGAAAATT	TAGTGTGTG	GCTAATCCTT	CAACATGGAA	ATGCCATAAA	AGTGACTTTG	ACAAAAAAA	AAGTATATAC	AGGTAGTAAA	CTCATCTACT
	CGACTTTTTAA	ATCACGACAC	CGATTAGGAA	GTTGTACCTT	TACGGTATTT	TCACTGAAAC	TGTTTTTTTT	TTCATATATG	TCCATCATTT	GAGTAGATGA
601	TCATTGACTT	TGTTTACAGC	ATGTGGAAGG	AGGAATATTT	ATTGCTAAAT	CGTAGTTTAA	CATTCAATAA	GTAATACTAT	TGAAATTCGA	CAAGATTGGC
	AGTAACTGAA	ACAAATGTCG	TACACCTTCC	TCCTTATAAA	TAACGATTTA	GCATCAAATT	GTAAGTTATT	CATTATGATA	ACTTTAAGCT	GTTCTAACCG
701	CGCATGGATG	AAAAAGAGGC	ATTTTGCTTT	GGGAGAATTA	GTTCAAATTA	GAAGTAAAA	AAAAAATTT	ACGAGGCAAA	AATGTCGGAT	TGAGATCGTA
	GCGTACCTAC	TTTTTCTCCG	TAAAACGAAA	CCCTCTTAAT	CAAGTTTAAT	CTTGACTTTT	TTTTTTGAAA	TGCTCCGTTT	TTACAGCCTA	ACTCTAGCAT
										ScaI
										~~~~~
801	AAAGTTCGCT	CGTCGTCTTT	TGCTTTGTGA	TTGTTTTTCAT	GGATACATCT	TGCTGGATAT	TTAAATTTTA	GTAATATGTA	TAAGATATTC	TATAAATGTT
	TTTCAAGCGA	GCAGCAGAAA	ACGAAACACT	AACAAAAGTA	CCTATGTAGA	ACGACCTATA	AATTTAAAAA	CATGATACAT	ATTCTATAAG	ATATTTACAA
901	TTATCACCCA	AACCTGTTAG	CGCCTTCTTA	ATTCTATTCA	ATCTGGCTTT	TGCTCTGAGA	CTACTTCTTG	GACTTTCACT	ACTTGTTAGT	TATACGGAAT
	AATAGTGGGT	TTGGACAATC	GCGGAAGAAT	TAAGATAAGT	TAGACCGAAA	ACGAGACTCT	GATGAAGAAC	CTGAAAGTGA	TGAACAATCA	ATATGCCTTA
										SacI EcoRI
										~~~~~
1001	TTGTGTAATT	AGAAGTGAAA	TAATCCTTTC	TATTAGTAAT	GCGAGCTCGA	ATTCGAGTCT	AACTCCTTAA	CCACTCGGAC	ATAGTGACTT	ATCTGACACT

AACACATTAA TCTTCACTTT ATTAGGAAAAG ATAATCATT A CGCTCGAGCT TAAGCTCAGA TTGAGGAATT GGTGAGCCTG TATCACTGAA TAGACTGTGA  
1101 CATTGTAAAT TAATATATAT ATAGGCATTT TGTTTAGTTA AAGGTA CTTA AGTAATTAGT ATAAACGAAC CAATTTTATA ATCAGGAAGT TAAGTGAATG  
GTAACATTTA ATTATATATA TATCCGTA AAA ACAAATCAAT TTCCATGAAT TCATTAATCA TATTTGCTTG GTTAAAATAT TAGTCCTTCA ATTCACCTTAC

XbaI  
~~~~~

1201 GTAGCACATG TCGTAAAAAT TGTGAATTTT TATTGAATAA TATTTTAAAT ACAAGCCTTT CTAGACTAGG TATACTCATA AATATGAGCA AAAGGATAGA  
CATCGTGTAC AGCATTTTTA AACTTAAAA ATA AACTTATT ATAAAAATTA TGTTCGGA AAA GATCTGATCC ATATGAGTAT TTATACTCGT TTCCTATCT

ScaI  
~~~~~

1301 GGAGATTACA TTGCATCTTC TACAAATTTA TTTATTGCC TTTACTGAAA AATTAAATAA TGAGTACTAA ATGATAAAAA GCGCTCAGTA CAAGAAAGAT  
CCTCTAATGT AACGTAGAAG ATGTTTAAAT AAATAACGGG AAATGACTTT TTAATTTATT ACTCATGATT TACTATTTTT CGCGAGTCAT GTTCTTTCTA

ScaI  
~~~~~

1401 TGCAAAAATA TTGCATTCTT CATGAATTAA AGTTGCATAT AAGGCATATT GAAAGTAATA GACTAAAAAC AGCAGTTAGC GAAAATTAAT AGAATTATAT  
ACGTTTTTAT AACGTAAGAA GACTTAAAT TCAACGTATA TTCCGTATAA CTTTCATTAT CATGATTTTG TCGTCAATCG CTTTTAATTA TCTTAATATA

1501 TCGCAAGACA ATTGTGACAT TAAATTA AAA AATTGTAGAA TTTTACTAT CCTCTTAAAC GCCATGAGCC TTTATAAAAA GGTAAATTA GTTTAACAT  
AGCGTTCTGT TAACACTGTA ATTTAATTTT TTAACATCTT AAAAATGATA GGAGAAATG CCGTACTCGG AAATATTTTT CCAATTTAAT CAAAATTGTA

XbaI  
~~~~~

1601 TCTTTTTTTG AGTAAGAGTT TACAATTTAT CAAAACCTGT GTTATTATAT TCATTAGTTT CAATTTATTA GCATCTAGAG AAAAATCAAT TGGCAGTTAC  
AGAAAAAAC TCATTCTCAA ATGTTAAATA GTTTTGGACA CAATAATATA AGTAATCAA GTTAAATAAT CGTAGATCTC TTTT TAGTTA ACCGTC AATG

1701 ATTGTTGGAA TTTATGAAGA AAAAGAATCT ACAACGGAGA ATAAGTTGCT GATCGCTTTC CCTAAAATTG TATATTTTGC TGAGCTTATT TTGACATTTT  
TAACAACCTT AAATACTTCT TTTTCTTAGA TGTTGCCCTCT TATTCACGA CTAGCGAAAG GGATTTTAAAC ATATAAACG ACTCGAATAA AACTGTAAAG

1801 GTTGAAGTTT TCTCTAATTC GCATTCATTT TAAGTAAAAC AATGAGAAAT AAAATTACAA AAAATACAAA TTAATAACA ATTTT TAGCT ATAATTATAG  
CAACTTCAAA AGAGATTAAG CGTAAGTAAA ATTCATTTTG TTTACTCTTA TTTTAAATGTT TTTTATGTTT AATTTTATGT TAAAAATCGA TATTAATATC

1901 ACGATGCCCT TGTATCCCAT TCTGTCTCGC TTGCCCTAC TTTTATCTT TTATATACCA TAATGAACGC TGCCGCTACT AACCATACCC CGATTTTACA  
TGCTACGGGA ACATAGGGTA AGACAGAGCG AACGGGGATG AAAAATAGAA AATATATGGT ATTACTTGCG ACGCGATGA TTGGTATGGG GCTAAAATGT

MluI  
~~~~~

2001 TTTCGGACTC CCAAGGACGT ACAAATAGA AACTATAGA AAAAAATAAT CAGAAAATAG CATGTCATCT CTTTGTA AAA CGCGTTTGCA AGAAGAAAGG  
AAAGCCTGAG GTTCTCTGCA TGTTTTATCT TTTGATATCT TTTTTTATA GTCTTTTATC GTACAGTAGA GAAACATTTT GCGCAAACGT TCTTCTTTCC

2101 AAACAATGGA GAAGAGATCA TCCATTTGTA TGTAATTTT AGTAAACTTG AAGAAATCAC TAACAACCTT TCTTACTTAG GGATTCATG CAAAACCTTG  
TTTGTACCT CTCTCTAGT AGGTAAACAT ACATTTAAAA TCATTTGAAC TTCTTTAGTG ATTGTTGAAG AGAATGAATC CCTAAGATAC GTTTTGG AAC

EcoRI

2201 TAAATCATCT GATGGAGGAC TCGATTTAAT GAATTGGAAG GTTGGAATTC ACTGGCCGTC GTTTACAAAC GTCGTGACTG GGAAAACCCT GGCGTTACCC  
ATTTAGTAGA CTACCTCCTG AGCTAAATTA CTTAACCTTC CAACCTTAAG TGACCGGCAG CAAAATGTTG CAGCACTGAC CCTTTTGGGA CCGCAATGGG

PvuII

PvuI

2301 AACTTAATCG CCTTGCAGCA CATCCCCCTT TCGCCAGCTG GCGTAATAGC GAAGAGGCCG GCACCGATCG CCCTTCCCAA CAGTTGCGCA GCCTGAATGG  
TTGAATTAGC GGAACGTCGT GTAGGGGGAA AGCGGTCGAC CGCATTATCG CTTCTCCGGG CGTGGCTAGC GGAAGGGTT GTCAACGCGT CGGACTTACC

NarI

2401 CGAATGGCGC CTGATGCGGT ATTTTCTCCT TACGCATCTG TCGGGTATTT CACACCGTAT GGTGCACTCT CAGTACAATC TGCTCTGATG CCGCATAGTT  
GCTTACCGCG GACTACGCCA TAAAAGAGGA ATGCGTAGAC ACGCCATAAA GTGTGGCATA CCACGTGAGA GTCATGTTAG ACGAGACTAC GGCGTATCAA

2501 AAGCCAGCCC CGACACCCGC CAACACCCGC TGACGCGCCC TGACGGGCTT GTCTGTCTCC GGCATCCGCT TACAGACAAG CTGTGACCGT CTCCGGGAGC  
TTCGGTCCGG GCTGTGGGCG GTTGTGGGCG ACTGCGCGGG ACTGCCCCGAA CAGACGAGGG CCGTAGGCGA ATGTCTGTTC GACACTGGCA GAGGCCCTCG

2601 TGCATGTGTC AGAGGTTTTT ACCGTCATCA CCGAAACGCG CGAGACGAAA GGGCCTCGTG ATACGCCTAT TTTTATAGGT TAATGTCATG ATAATAATGG  
ACGTACACAG TCTCCAAAAG TGGCAGTAGT GGCTTTGCGC GCTCTGCTTT CCCGGAGCAC TATGCGGATA AAAATATCCA ATTACAGTAC TATTATTACC

2701 TTTCTTAGAC GTCAGGTGGC ACTTTTCGGG GAAATGTGCG CGGAACCCCT ATTTGTTTAT TTTTCTAAAT ACATTCAAAT ATGTATCCGC TCATGAGACA  
AAAGAATCTG CAGTCCACCG TGAAAAGCCC CTTTACACGC GCCTTGGGGA TAAACAAATA AAAAGATTTA TGTAAGTTTA TACATAGGCG AGTACTCTGT

2801 ATAACCCTGA TAAATGCTTC AATAATATTG AAAAAAGGAAG AGTATGAGTA TTCAACATTT CCGTGTGCGC CTTATTCCCT TTTTTCGCGC ATTTTGCCTT  
TATTGGGACT ATTTACGAAG TTATTATAAC TTTTTCCTTC TCATACTCAT AAGTTGTAAA GGCACAGCGG GAATAAGGGA AAAAAAGCGC TAAAAAGGAA

2901 CCTGTTTTTG CTCACCCAGA AACGCTGGTG AAAGTAAAAG ATGCTGAAGA TCAGTTGGGT GCACGAGTGG GTTACATCGA ACTGGATCTC AACAGCGGTA  
GGACAAAAC GAGTGGGTCT TTGCGACCAC TTTTATTTTC TACGACTTCT AGTCAACCCA CGTGTCTACC CAATGTAGCT TGACCTAGAG TTGTGCGCAT

3001 AGATCCTTGA GAGTTTTTCG CCCGAAGAAC GTTTTCCAAT GATGAGCACT TTTAAAGTTC TGCTATGTGG CGCGGTATTA TCCCGTATTG ACGCCGGGCA  
TCTAGGAACT CTCAAAAAGCG GGGCTTCTTG CAAAAGGTTA CTACTCGTGA AAATTTCAAG ACGATACACC GCGCCATAAT AGGCATAAC TGCGGCCCGT

ScaI

3101 AGAGCAACTC GGTCGCCGCA TACACTATTC TCAGAATGAC TTGGTTGAGT ACTCACCAGT CACAGAAAAG CATCTTACGG ATGGCATGAC AGTAAGAGAA  
TCTCGTTGAG CCAGCGGCGT ATGTGATAAG AGTCTTACTG AACCAACTCA TGAGTGGTCA GTGTCTTTTC GTAGAATGCC TACCGTACTG TCATTCTCTT

PvuI

3201 TTATGCAGTG CTGCCATAAC CATGAGTGAT AACACTGCGG CCAACTTACT TCTGACAACG ATCGGAGGAC CGAAGGAGCT AACCGTTTTT TTGCACAACA  
AATACGTAC GACGGTATTG GTACTACTA TTGTGACGCC GGTGTAATGA AGACTGTTGC TAGCCTCCTG GCTTCCTCGA TTGCGGAAAA AACGTGTTGT

3301 TGGGGATCA TGTAACCTCG CTTGATCGTT GGAACCGGA GCTGAATGAA GCCATACCAA ACGACGAGCG TGACACCACG ATGCTGTAG CAATGGCAAC  
ACCCCTAGT ACATTGAGCG GAACTAGCAA CCCTTGGCCT CGACTTACTT CGGTATGGTT TGCTGCTCGC ACTGTGGTGC TACGGACATC GTTACCGTTG

3401 AACGTTGCGC AACTATTAA CTGGCGAAT ACTTACTCTA GCTTCCCGC AACAAATAAT AGACTGGATG GAGGCGGATA AAGTTGCAGG ACCACTTCTG

TTGCAACGCG TTTGATAAATT GACCGCTTGA TGAATGAGAT CGAAGGGCCG TTGTTAATTA TCTGACCTAC CTCCGCCTAT TTCAACGTCC TGGTGAAGAC  
 3501 CGCTCGGCC TCCGGCTGG CTGGTTTATT GCTGATAAAT CTGGAGCCGG TGAGCGTGGG TCTCGCGGTA TCATTGCAGC ACTGGGGCCA GATGGTAAGC  
 GCGAGCCGGG AAGGCCGACC GACCAAATAA CGACTATTTA GACCTCGGCC ACTCGCACCC AGAGCGCCAT AGTAACGTCC TGACCCCGGT CTACCATTCC  
 3601 CCTCCCGTAT CGTAGTTATC TACACGACGG GGAGTCAGGC AACTATGGAT GAACGAAATA GACAGATCGC TGAGATAGGT GCCTCACTGA TTAAGCATTG  
 GGAGGGCATA GCATCAATAG ATGTGCTGCC CCTCAGTCCG TTGATACCTA CTTGCTTTAT CTGTCTAGCG ACTCTATCCA CGGAGTGACT AATTTCGTAAC  
 3701 GTAACGTGCA GACCAAGTTT ACTCATATAT ACTTTAGATT GATTTAAAAC TTCATTTTTA ATTTAAAAGG ATCTAGGTGA AGATCCTTTT TGATAATCTC  
 CATTGACAGT CTGGTTCAAA TGAGTATATA TGAATCTAA CTAATTTTG AAGTAAAAAT TAAATTTTCC TAGATCCACT TCTAGGAAAA ACTATTAGAG  
 3801 ATGACCAAAA TCCCTTAACG TGAGTTTTTCG TTCCACTGAG CGTCAGACCC CGTAGAAAAG ATCAAAGGAT CTTCTTGAGA TCCTTTTTTT CTGCGCGTAA  
 TACTGGTTTT AGGAATTGC ACTCAAAGC AAGGTGACTC GCAGTCTGGG GCATCTTTTC TAGTTTCCCTA GAAGAACTCT AGGAAAAAAA GACGCGCATT  
 3901 TCTGCTGCTT GCAAACAAAA AAACCACCGC TACCAGCGGT GGTGTTGTTG CCGGATCAAG AGCTACCAAC TCTTTTCCG AAGGTAAGT GCTTCAGCAG  
 AGACGACGAA CGTTTGTTTT TTTGGTGGCG ATGGTCCGCA CCAAACAAA GGCCTAGTTC TCGATGGTTG AGAAAAAGGC TTCCATTGAC CGAAGTCTGC  
 4001 AGCGCAGATA CCAAATACTG TCCTTCTAGT GTAGCCGTAG TTAGGCCACC ACTTCAAGAA CTCTGTAGCA CCGCTACAT ACCTCGCTCT GCTAATCCTG  
 TCGCGTCTAT GGTTTATGAC AGGAAGATCA CATCGGCATC AATCCGGTGG TGAAGTTCTT GAGACATCGT GCGCGATGTA TGGAGCGAGA CGATTAGGAC  
 4101 TTACCAGTGG CTGCTGCCAG TGGCGATAAG TCGTGTCTTA CCGGGTTGGA CTCAAGACGA TAGTTACCGG ATAAGGCGCA GCGGTCGGGC TGAACGGGGG  
 AATGGTCACC GACGACGGTC ACCGCTATTC AGCACAGAAT GGCCCAACCT GAGTTCTGCT ATCAATGGCC TATTCGCGT GCGCAGCCCG ACTTGCCCCC  
 4201 GTTCGTGCAC ACAGCCAGC TTGGAGCGAA CGACCTACAC CGAACTGAGA TACCTACAGC GTGAGCTATG AGAAAGCGCC ACGCTTCCCG AAGGGAGAAA  
 CAAGCACGTG TGTCGGGTCG AACCTCGCTT GCTGGATGTG GCTTGACTCT ATGGATGTCG CACTCGATACT TCTTTCGCGG TCGGAAGGGC TTCCCTCTTT  
 4301 GCGGACAGG TATCCGGTAA GCGGCAGGGT CGGAACAGGA GAGCGCACGA GGGAGCTTCC AGGGGAAAAC GCCTGGTATC TTTATAGTCC TGTCGGGTTT  
 CCGCCTGTCC ATAGGCCATT CGCCGTCCCA GCCTTGTCTT CTCGCGTGT CCCTCGAAGG TCCCCCTTG CCGACCATAG AAATATCAGG ACAGCCCAAA  
 4401 CGCCACCTCT GACTTGAGCG TCGATTTTTG TGATGCTCGT CAGGGGGGCG GAGCCTATGG AAAAAAGCCA GCAACGCGGC CTTTTTACGG TTCTTGCCCT  
 GCGGTGGAGA CTGAACTCGC AGCTAAAAAC ACTACGAGCA GTCCCCCGC CTCGATACC TTTTTGCGGT CGTTGCGCCG GAAAAATGCC AAGGACCGGA  
 4501 TTTGCTGGCC TTTTGTCTAC ATGTTCTTTC CTGCGTTATC CCCTGATTCT GTGGATAACC GTATTACCGC CTTTGAGTGA GCTGATACCG CTCGCCGAG  
 AAACGACCGG AAAACGAGTG TACAAGAAAG GACGCAATAG GGGACTAAGA CACCTATTGG CATAATGGCG GAAACTCACT CGACTATGGC GAGCGGCGTC  
  
 PvuII  
 ~~~~  
 4601 CCGAACGACC GAGCGCAGCG AGTCAGTGAG CGAGGAAGCG GAAGAGCGCC CAATACGCAA ACCGCCTCTC CCCGCGGTT GCGCGATTCA TTAATGCAGC  
 GGCTTGCTGG CTCGCGTCGC TCAGTCACTC GCTCCTTCGC CTTCTCGCGG GTTATGCGTT TGGCGGAGAG GGGCGCGCAA CCGGCTAAGT AATTACGTCC  
  
 PvuII  
 ~~  
 4701 TGGCACGACA GGTTCCTCCGA CTGGAAAGCG GGCAGTGAGC GCAACGCAAT TAATGTGAGT TAGCTCACTC ATTAGGCACC CCAGGCTTTA CACTTTATGC  
 ACCGTGCTGT CCAAAGGGCT GACCTTTCGC CCGTCACTCG CGTTGCGTTA ATTACTCTCA ATCGAGTGAG TAATCCGTGG GTTCCGAAAT GTGAAATACG

HindIII

4801 TTCCGGCTCG TATGTTGTGT GGAATTGTGA GCGGATAACA ATTTACACA GGAAACAGCT ATGACCATGA TTACGCCAAG CTTTACGTCG TAAGGCCGTT  
 AAGGCCGAGC ATACAACACA CCTTAACACT CGCCTATTGT TAAAGTGTGT CCTTTGTCGA TACTGGTACT AATGCGGTTT GAAATGCAGC ATTCGGGCAA

4901 TCTGACAGAG TAAAATTCTT GAGGGAACCT TCACCATTAT GGGAAATGGT TCAAGAAGGT ATTGACTTAA ACTCCATCAA ATGGTCAGGT CATTGAGTGT  
 AGACTGTCTC ATTTTAAGAA CTCCTTGAA AGTGGTAATA CCCTTTACCA AGTTCCTTCCA TAACTGAATT TGAGGTAGTT TACCAGTCCA GTAACCTACA

5001 TTTTATTTG TTGTATTTTT TTTTTTTTAG AGAAAATCCT CCAATATCAA ATTAGGAATC GTAGTTTCAT GATTTTCTGT TACACCTAAC TTTTGTGTG  
 AAAAAATAAC AACATAAAAA AAAAAAATC TCTTTTAGGA GGTATAGTT TAATCCTTAG CATCAAAGTA CTAAAAGACA ATGTGGATTG AAAAAACAC

5101 GTGCCCTCCT CTTGTCAAT ATTAATGTTA AAGTGAATF CTTTTCTCTT ATCACGTGA GCCATTAGTA TCAATTTGCT TACCTGTATT CCTTACTAT  
 CACGGGAGGA GGAACAGTTA TAATTACAAT TTCACGTTAA GAAAAAGGAA TAGTGCAACT CGGTAATCAT AGTTAAACGA ATGGACATAA GGAAATGATA

5201 CCTCCTTTTT CTCTTCTTG ATAAATGTAT GTAGATTGCG TATATAGTTF CGTCTACCTT ATGAACATAT TCCATTTTGT AATTTCTGTT CGTTTCTATT  
 GGAGGAAAAA GAGGAAGAAC TATTTACATA CATCTAACGC ATATATCAA GCAGATGGGA TACTTGTATA AGGTAAAACA TTAAAGCACA GCAAAGATAA

5301 ATGAATTTCA TTTATAAAGT TTATGTACAA ATATCATAAA AAAAGAGAAT CTTTTTAAGC AAGGATTTTC TTAACCTCTT CGGCGACAGC ATCACCGACT  
 TACTTAAAGT AAATATTTCA AATACATGTT TATAGTATTT TTTTCTCTTA GAAAAATTCG TTCCTAAAAG AATTGAAGAA GCCCTGTCTG TAGTGGCTGA

5401 TCGGTGGTAC TGTGGAACC ACCTAAATCA CCAGTTCCTGA TACCTGCATC CAAAACCTTT TTAACCTGCAT CTTCAATGGC CTTACCTTCT TCAGGCAAGT  
 AGCCACCATG ACAACCTTGG TGGATTTAGT GGTCAAGACT ATGGACGTAG GTTTTGGAAA AATTGACGTA GAAGTTACCG GAATGGAAGA AGTCCGTTCA

5501 TCAATGACAA TTTCAACATC ATTGCAGCAG ACAAGATAGT GCGGATAGGG TCAACCTTAT TCTTTGGCAA ATCTGGAGCA GAACCGTGGC ATGGTTCGTA  
 AGTFACTGTT AAAGTTGTAG TAACGTCTGTC TGTTCTATCA CCGCTATCCC AGTTGGAATA AGAAAACCGTT TAGACCTCGT CTTGGCACCG TACCAAGCAT

EcoRV

~~~~

5601 CAAACCAAAT GCGGTGTTCT TGTCTGGCAA AGAGGCCAAG GACGCAGATG GCAACAAACC CAAGGAACCT GGGATAACGG AGGCTTCATC GGAGATGATA  
 GTTTGGTTTA CGCCACAAGA ACAGACCGTT TCTCCGGTTC CTGCGTCTAC CGTTGTTTGG GTTCCTTGGA CCCTATTGCC TCCGAAGTAG CCTCTACTAT

EcoRV

~~

5701 TCACCAAACA TGTGCTGGT GATTATAATA CCATTTAGGT GGGTTGGGTT CTTAACTAGG ATCATGGCGG CAGAATCAAT CAATTGATGT TGAACCTTCA  
 AGTGGTTTGT ACAACGACCA CTAATATTAT GGTAAATCCA CCCAACCCTA GAATTGATCC TAGTACCGCC GTCTTAGTTA GTTAACTACA ACTTGGAAGT

5801 ATGTAGGGTA TTCGTTCTTG ATGGTTTCCT CCACAGTTTT TCTCCATAAT CTTGAAGAGG CAAAAACATT AGCTTTATCC AAGGACCAAA TAGGCAATGG  
 TACATCCCAT AAGCAAGAAC TACCAAAGGA GGTGTCAAAA AGAGGTATTA GAACCTTCTC GGTTTTGTAA TCGAAATAGG TTCCTGGTTT ATCCGTTACC

5901 TGGCTCATGT TGTAGGGCCA TGAAAGCGGC CATTCTTGTG ATTCTTTGCA CTTCTGGAAC GGTGTATTGT TCACTATCCC AAGCGACACC ATCACCATCG  
 ACCGAGTACA ACATCCCGGT ACTTTCGCCG GTAAGAACAC TAAGAAACGT GAAGACCTTG CCACATAACA AGTGATAGGG TTCCTGTGG TAGTGGTAGC

6001 TCTTCCTTTC TCTTACAAA GTAAATACCT CCCACTAATT CTCTGACAAC AACGAAGTCA GTACCTTTAG CAAATTGTGG CTTGATTGGA GATAAGTCTA  
 AGAAGGAAAG AGAATGGTTT CATTTATGGA GGGTGATTAA GAGACTGTGT TTGCTCAGT CATGGAAATC GTTTAACACC GAACCTAACCT CTATTCAGAT

KpnI

~~~~

6101 AAAGAGAGTC GGATGCAAAG TTACATGGTC TTAAGTTGGC GTACAATTGA AGTTCTTTAC GGATTTTTFAG TAAACCTTGT TCAGGTCTAA CACTACCGGT  
TTTCTCTCAG CCTACGTTTC AATGTACCAG AATTCAACCG CATGTTAACT TCAAGAAATG CCTAAAAATC ATTTGGAACA AGTCCAGATT GTGATGGCCA

KpnI

~~~

ClaI

~~~~~

6201 ACCCCATTTA GGACCACCCA CAGCACCTAA CAAAACGGCA TCAACCTTCT TGGAGGCTTC CAGCGCCTCA TCTGGAAGTG GGACACCTGT AGCATCGATA  
TGGGGTAAAT CCTGGTGGGT GTCGTGGATT GTTTTGCCGT AGTTGGAAGA ACCTCCGAAG GTCGCGGAGT AGACCTTCAC CCTGTGGACA TCGTAGCTAT

6301 GCAGCACCAC CAATTAATG ATTTTCGAAA TCGAACTTGA CATTGGAACG AACATCAGAA ATAGCTTTAA GAACCTTAAT GGCTTCGGCT GTGATTTCTT  
CGTCGTGGTG GTTAATTTAC TAAAAGCTTT AGCTTGAAC GTAACCTTGC TTGTAGTCTT TATCGAAATT CTTGGAATTA CCGAAGCCGA CACTAAAGAA

6401 GACCAACGTG GTCACCTGGC AAAACGACGA TCTTCTTAGG GGCAGACATT AGAATGGTAT ATCCTTGAAA TATATATATA TATTGCTGAA ATGTAAGG  
CTGGTTGCAC CAGTGGACCG TTTTGCTGCT AGAAGAATCC CCGTCTGTAA TCTTACCATA TAGGAACTTT ATATATATAT ATAACGACTT TACATTTTCC

6501 TAAGAAAAGT TAGAAAAGTAA GACGATTGCT AACCACCTAT TGGAAAAAAC AATAGGTCCT TAAATAATAT TGTCAACTTC AAGTATTGTG ATGCAAGCAT  
ATTCTTTTCA ATCTTTCATT CTGCTAACGA TTGGTGGATA ACCTTTTTTG TTATCCAGGA ATTTATTATA ACAGTTGAAG TTCATAACAC TACGTTTCGA

6601 TTAGTCATGA ACGCTTCTCT ATTCTATATG AAAAGCCGGT TCCGGCCTCT CACCTTTCCT TTTTCTCCA ATTTTTCAGT TGAAAAAGGT ATATGCGTCA  
AATCAGTACT TGCGAAGAGA TAAGATATAC TTTTCCGGCA AGGCCGGAGA GTGGAAAGGA AAAAGAGGGT TAAAAAGTCA ACTTTTTTCCA TATACGCAGT

6701 GCGGACCTCT GAAATTAACA AAAAATTTCC AGTCATCGAA TTTGATTCTG TGCGATAGCG CCCCTGTGTG TTCTCGTTAT GTTGAGGAAA AAAATAATGG  
CCGCTGGAGA CTTTAATTGT TTTTAAAGG TCAGTAGCTT AAATAAGAC ACGCTATCGC GGGGACACAC AAGAGCAATA CAACTCCTTT TTTTATTACC

NarI

~~~~~

6801 TTGCTAAGAG ATTCGAAGTC TTGCATCTTA CGATACCTGA GTATTCCCAC AGTTAACTGC GGTCAAGATA TTTCTTGAAT CAGGCGCCTT AGACCGCTCG  
AACGATTCTC TAAGCTTGAG AACGTAGAAT GCTATGGACT CATAAGGGTG TCAATTGACG CCAGTTCATAT AAAGAACTTA GTCCGCGGAA TCTGGCGAGC

6901 GCCAAACAAC CAATTACTTG TTGAGAAATA GAGTATAATT ATCCTATAAA TATAACGTTT TTGAACACAC ATGAACAAGG AAGTACAGGA CAATTGATTT  
CGGTTTGTTG GTTAATGAAC AACTCTTTAT CTCATATTAA TAGGATATTT ATATTGCAAA AACTTGTGTG TACTTGTTCC TTCATGTCTT GTTAACTAAA

XhoI

~~~~~

HindIII

~~~~~

7001 TGAAGAGAAT GTGGATTTTG ATGTAATTGT TGGGATTCCA TTTTAAATAA GGCAATAATA TTAGGTATGT GGATATACTA GAAGTTCTCC TCGAGAAGCT  
ACTTCTCTTA CACCTAAAAC TACATTAACA ACCCTAAGGT AAAAATTATT CCGTTATTAT AATCCATACA CCTATATGAT CTTCAAGAGG AGCTCTTCGA

HindIII PstI

~

~~~~~

7101 TGCATGCCTG CAGGTGCGAGT CAGAAAATTA TCGCCATAAA AGACAGAATA AGTCATCAGC GGTTGTTTCA TTTCTATAT TTTTTTTTTA TTTTTTTATT  
ACGTACGGAC GTCCAGCTCA GTCTTTTAAT AGCGGTATTT TCTGTCTTAT TCAGTAGTCG CCAACAAAGT AAAGGATATA AAAAAAAAT AAAAAAATAA

7201 TTTTAATAAG GGAAAATTTA ACGTCTAAGG ATACAGAAGA TTGTTAGCAC ATTAAGTAA TAAAGGCTTA AGTAGTAAGT GCCTTAGCAT GTTATTGTAT  
AAAATTATTC CCTTTTAAAT TGCAGATTCC TATGTCTTCT AACAAATCGTG TAATTTTCATT ATTTCCGAAT TCATCATTCA CGGAATCGTA CAATAACATA

7301 TTCAAAGGAC ATAATCTAAA ATAATAACAA TATCATTTCT CACAAGTTAT TCAATTTTCT TTTTTTTTTC TAATAATATC AAGAATGTAT TATTTGTTTG  
AAGTTTCCTG TATTAGATTT TATTATTGTT ATAGTAAAGA GTGTTCAATA AGTTAAAAGA AAAAAAAAAA ATTATTATAG TTCTTACATA ATAAACAAAC

7401 ACATAAGTCA ACTAATTTAT TTAATATGCT GGATTAATCT TGCAGACATG TAAATTAACA GTTTTAGTCA AATAACGTTG AAGTTTCAAT GAACTCAAAT  
TGTATTCAGT TGATTAAATA AATTATACGA CCTAATTAGA ACGTCTGTAC ATTTAATTGT CAAAATCAGT TTATTGCAAC TTCAAAGTTA CTTGAGTTTA

7501 AATTTCTCTT TTTTTTTATA TAACATGTCT AATCTGATTT ATATTTTCCG CAGGATCAAC TGAAGTTATG ACATTTGGAT TGGATCACTT ATAACCTTGG  
TTAAAGAGAA AAAAAAATAT ATTGTACAGA TTAGACTAAA TATAAAAGGC GTCCTAGTTG ACTTCAATAC TGTAAACCTA ACCTAGTGAA TATTGGAACC

7601 TCGCCAAATA ATACAAAAAT CAGCGTTATA AAACAAAGAG GTTTTTGTTA AGAAATTAAT CCTCTTCTT GATAAGAAAG TTGAACCGAA ATGCAGATA  
AGCGGTTTAT TATGTTTTTA GTCGCAATAT TTTGTTTCTC CAAAAACAAT TCTTTAATTA GGAGAAAGAA CTATCTTTC AACTTGGCTT TAACGTCTAT

7701 CTGATATATG AAAATAATAC CCACAATTTT GGAATAGCG CAAGCCTCAA TTTAAACAAT AGGTGAGGAC ACATGATAAT GACCTCAATG ATGTTAGAA  
GACTATATAC TTTTATTATG GGTGTTAAAA CCCTTATCGC GTTCGGAGTT AAATTTGTTA TCCACTCCTG TGTACTATTA CTGGAGTTAC TAACAATCTT

7801 GAAAAGAGCC TCATTACAAA ATCGAAAAAT GAATGGTTGG GTACAAGTTT CCAAACATG GTAAAGTGGA CTTTGCGTAT GAGACGTAAA TAGAAAAAAA  
CTTTTCTCGG AGTAATGTTT TAGCTTTTTA CTTACCAACC CATGTTCAAA GGTTTTGTAC CATTTACCT GAAACGCATA CTCTGCATTT ATCTTTTTTT

XbaI

~~~~~

7901 CACTTGTTAT ATGTTTTCTA GAATTATTGT TGTCTCTTTA TGGTTGGATG ATGCAAATA GTAATTTCTG TTAGTTGCTG TAAAAACCA CGAGACAAAT  
GTGAACAATA TACAAAAGAT CTTAATAACA ACAGAGAAAT ACCAACCTAC TACGTTTAT CATTTAAAGCC AATCAACGAC ATTTTGTGGT GCTCTGTTTA

8001 AGATATGGAT ATTTATTAATA TCAGGAAAAA CGTAACCTCTC GGCTACTGGA TGGTTCAGTC ACCCAACGAT TACTGGGGAG AGAAAAACAG GCAAAAAGCAT  
TCTATACCTA TAAATAATTT AGTCCTTTTT GCATTGAGAG CCGATGACCT ACCAAGTCAG TGGGTTGCTA ATGACCCCTC TCTTTTGTCC CGTTTTCGTA

8101 AGCTTAAAGG AATCCGATTG TCATTCCGCA ATGTGCAGCG AAACCTAAAA CCGGATAATG GACCTGTTAA TCGAAACAAT TGAAGATATA TAAAGGAAGA  
TCGAATTTCC TTAGGCTAAC AGTAAGCCGT TACACGTGCG TTTGATTTTT GGCCTATTAC CTGGACAATT AGCTTTGTTA ACTTCTATAT ATTTCTTCTT

PacI

~~~~~

8201 GGAATCCTGG CATATCATCA ATTGAATAAG TTGAATTAAT TATTTCAATC TCATTCTCAC TTTCTGACTT ATAGTCGCTT TGTTAACATA TTAATTAATG  
CCTTAGGACC GTATAGTAGT TAACTTATTC AACTTAATTA ATAAAGTTAG AGTAAGAGTG AAAGACTGAA TATCAGCGAA ACAATTGTAT AATTAATTAC

8301 AAAGCTGATG CGCAACAAAA TAACTTCAAC AAAGATCAAC AAAGCGCCTT CTATGAAATC TTGAACATGC CTAACCTAAA CGAAGCGCAA CGTAACGGCT  
TTTCGACTAC GCGTTGTTTT ATTGAAGTTG TTTCTAGTTG TTTCCGGGAA GATACTTTAG AACTTGTACG GATTGAATTT GCTTCGCGTT GCATTGCCGA

8401 TCATTCAAAG TCTTAAAGAC GACCCAAGCC AAAGCACTAA CGTTTTAGGT GAAGCTAAAA AATTAACGA ATCTCAAGCA CCGAAAGCTG ATAACAATTT  
AGTAAGTTTC AGAATTTCTG CTGGGTTCCG TTTCTGTGATT GCAAAATCCA CTTCGATTTT TTAATTTGCT TAGAGTTCGT GGCTTTTCGAC TATTGTAAAA

HindIII

~~~~~

8501 CAACAAAGAA CAACAAATG CTTTCTATGA AATCTTGAAT ATGCCTAACT TAAACGAAGA ACAACGCAAT GGTTTCATCC AAAGCTTAAA AGATGACCCA  
GTTGTTTCTT GTTGTTTTAC GAAAGATACT TTAGAACCTA TACGGATTGA ATTTGCTTCT TGTTCGTTA CCAAAGTAGG TTTCGAATTT TCTACTGGGT

8601 AGCCAAAGTG CTAACCTATT GTCAGAAGCT AAAAAGTTAA ATGAATCTCA AGCACCGAAA GCGGATAACA AATCAACAA AGAATCTAGT ACCCCAACTA

TCGGTTTCAC GATTGGATAA CAGTCTTCGA TTTTCAATT TACTTAGAGT TCGTGGCTTT CGCCTATTGT TTAAGTTGTT TCTTAGATCA TGGGGTTGAT

XhoI

~~~~~

8701 CTGCTTCTGA AAATCTATAT TTTCAAGGTG AACTAAAAAC TGCTGCTTTG GCTCAACATG CGCTCGAGAA GATGAAGCGA CGATGGAAAA AGAATTTTCAT  
GACGAAGACT TTTAGATATA AAAGTTCCAC TTGATTTTTG ACGACGAAAC CGAGTTGTAC GCGAGCTCTT CTACTIONGCT GCTACCTTTT TCTTAAAGTA

AscI

~~

8801 AGCCGTCTCA GCAGCCAACC GCTTTAAGAA AATCTCATCC TCCGGGGCAC TTGATTATGG  
TCGGCAGAGT CGTCGGTTGG CGAAATTCTT TTAGAGTAGG AGGCCCCGTG AACTAATACC