



PvuI (7)
SgfI (6) 1 GGATCTGGATCGCTCCGGTGCCCGTCAGTGGCAGAGCGCACATCGCCACAGTCCCGGAGAAGTTGGGGGAGGGGTGGCAATTGAACGGGTGCCTA **MfeI (82)**
101 GAGAAGGTGGCGCGGGTAAACTGGAAAGTGATGTCGTGTAAGTGGCTCCGCCTTTTCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC

Psp1406I (203) 201 GTGAACGTTCTTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGGCTCGCATCTCTCTTACACGGCCCGCCGCCCTACCTGAGGCC
HindIII (245)
PvuII (239) 301 GCCATCCACGCGGTTGAGTGCAGTCTGCCGCTCCCGCCTGTGGTGCCTCTGAAGTGCCTCCGCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC
Bsu36I (291)

NgoMIV (441)
NaeI (441) 401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCCTGCTTCTCAACTCTACGTCTTTGTTTCGTTT

NcoI (560)
BstEII (555)
KasI (535) 501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGCGCCCTAGCTGAGATCACCGGTACCCATGGGCTGTGTGCAATGTAAGGATAAAGAAGCAGCGAA
AgeI (552) 1 M G C V Q C K D K E A A K
601 ACTGACAGAGGAGAGGGACGGCAGCCTGAACAGAGCTCTGGTACCGCTATGGCACAGACCCACCCCTCAGCACTACCCAGCTTCGGCGTGACCTCC
13 L T E E R D G S L N Q S S G Y R Y G T D P T P Q H Y P S F G V T S

PvuII (725) 701 ATCCCGAACTACAACAACCTCCACGCAGCTGGGGCCAGGGACTCACCGTCTTTGGGGGTGTAAGTCTCTCTCACACTGGGACCCCTACGCACGAGAG
47 I P N Y N N F H A A G G Q G L T V F G G V N S S S H T G T L R T R
SandI (781)
EcoO109I (781) 801 GAGGGACAGGAGTGACACTGTTTGTGGCGCTTATGACTATGAAGCACGGACGGAAGATGACCTGAGTTTTCACAAAGGAGAAAAATTTCAAATTTGAA
80 G G T G V T L F V A L Y D Y E A R T E D D L S F H K G E K F Q I L N
SspI (891)

BsrBI (927) 901 CAGCTCGGAAGGAGACTGGTGGGAAGCCGCTCCTTGACAACCGGGGAACTGTTTACATTCAGCAATTACGTGGCTCCAGTTGACTCCATCCAGGCA
113 S S E G D W W E A R S L T T G E T G Y I P S N Y V A P V D S I Q A
1001 GAAGAGTGGTACTTTGAAAACTTGCCCGCAAAGATGCTGAGAGACAGCTCTGTCTTTGAAACCCAAAGAGTACCTTTCTATCCCGGAGAGCGAAA
147 E E W Y F G K L G R K D A E R Q L L S F G N P R G T F L I R E S E
XcmI (962)

Tth111I (1153) 1101 CCACAAAGGTGCCTACTCACTTCCATCCGTGATTGGGATGATATGAAAGGGGACCAGCTCAAACATTATAAAATCCGCAAGCTTGACAATGGTGGATA
180 T T K G A Y S L S I R D W D D M K G D H V K H Y K I R K L D N G G Y
HindIII (1180)

Bsp120I (1214) 1201 CTATATCACAACGCGGGCCAGTTTGAACACTTCAGCAACTGGTACAGCATTACTCAGAGAGAGCCGAGGTCTCTGCTGCCGCTAGTAGTCCCTGT
213 Y I T T R A Q F E T L Q Q L V Q H Y S E R A A G L C C R L V V P C
XhoI (1361)

1301 CACAAAGGATGCCAAGGCTTACCGATCTGTCTGTCAAACCAAAGATGTCTGGGAAATCCCTCGAGAATCCCTGCAGTTGATCAAGAGACTGGGAAATG
247 H K G M P R L T D L S V K T K D V W E I P R E S L Q L I K R L G N
BspEI (1482)

1401 GGCAGTTTGGGGAAGTATGGATGGGTACCTGGAATGAAATACAAAAGTAGCCATAAAGACCCCTAAGCCAGGCACCATGTCTCCGGAGTCCCTCCTGGA
280 G Q F G E V W M G T W N G N T K V A I K T L K P G T M S P E S F L E
BspHI [m] (1509)

1501 GGAGGCGCAGATCATGAAGAAGCTGAAGCATGACAAGCTGGTGCAGCTCTACGCGGTGCTGTCTGAGGAGCCATTTACATCGTCACGGAGTACATGAGC
313 E A Q I M K K L K H D K L V Q L Y A V V S E E P I Y I V T E Y M S
1601 AAAGGAAGTTGCTTACTTCTAAAAGATGGTGAAGGAAGAGCTCTGAAGTTGCCAAACCTTGTGGACATGGCGGCACAGTTGCTGCAGGAATGGCTT
347 K G S L L D F L K D G E G R A L K L P N L V D M A A Q V A A G M A
BglII (1728)

1701 ACATCGAGCGCATGAATTATATCCACAGAGATCTGCGATCAGCAAACATTCTAGTGGGAATGGACTAATTTGCAAGATTGCTGACTTTGGATTGGCTCG
380 Y I E R M N Y I H R D L R S A N I L V G N G L I C K I A D F G L A R
BbsI (1807)

1801 GTTGATTGAAGACAATGAATACACAGCAAGACAAGGTGCGAAGTTTCCATTAAGTGGACAGCCCCGAGCGCCCTGTATGGAAGTTTCAATCAAG
413 L I E D N E Y T A R Q G A K F P I K W T A P E A A L Y G R F T I K
BstEII (1938)

1901 TCTGACGTATGGTCTTTTGAATCTTACTCACAGAGCTGGTCAACAAAGGAGTGCATACCCAGGCATGAACAACCGGGAGGTGCTGGAGCAGGTGG
447 S D V W S F G I L L T E L V T K G R V P Y P G M N N R E V L E Q V
BspEI (2079)

2001 AGAGAGGCTATAGGATGCCCTGCCACAGGACTGCCCGATCTCCCTGCACGAGCTCATGATCCACTGTGAAAAAGGATCCGGAAAGAGCGCCCGACCTT
480 E R G Y R M P C P Q D C P I S L H E L M I H C W K K D P E E R P T F
BspHI [m] (2054) **BamHI (2076)**

ScaI (2102) 2101 CGAGTACTTGCAGGGCTTCTGGAGGACTACTTTACGGCCACAGAGCCCCAGTATCAGCCCGGTGAAAACCTGTGAGAGCCTGCGCTTCAGACGCTAGCT
513 E Y L Q G F L E D Y F T A T E P Q Y Q P G E N L •
2201 GCCCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACCTAGAATGCAGTGAATAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTT

HpaI (2331) **MfeI (2342)** 2301 ATTTGTAACCATTATAAGCTGCAATAAAACAAGTTAACAACAACAATTGCATTCTTTTATGTTTCAGGTTTCAGGGGGAGGTGGGAGGTTTTTAAAGC

EcoRI (2427)
 2401 AAGTAAAACCTCTACAAATGTGGTATGGAATCTAAAATACAGCATAGCAAACCTTTAACCTCCAATCAAGCCTCTACTTGAATCCTTTCTGAGGGAT
 2501 GAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTTGCAGCCTCACCTTCTTTCATGGAGTTAAGATATAGTGATTTTCCCAAGG

2601 TTTGAACTAGCTCTTCATTTCTTTATGTTTTAAATGCACTGACCTCCACATTCCCTTTTATGAAAAATTTCAGAAATAATTTAAATACATCATTGCAA
 SspI (2666) SwaI (2680)

2701 TGAATAAATGTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTCATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAAGGAACCTTTA
 EcoO109I (2741)

2801 ATAGAAATTGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCCTGGTACTTGGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCCATTC
 141 • N R T Y K L P I L E E I T T K V L K G N
 2901 ATCTCAATGAGCACAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTGACATGCCACAGGGGCTGACCACCTGATGGATCTGTCCACCTCATCAG
 119 M E I L V F C D P A Y D S I L E R C M G C P S V V R I S R D V E D S
 3001 AGTAGGGTGCCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCCGTTGCTCACAGCAGACCAATGGCAATGGCTTCCAGCAGACAGTGGCCTGCC
 86 Y P H R V A V I T D F D K Q G N S V A S G I A I A E A C V T V R G

3101 AATGTAGGCCTCAATGTGGACAGCAGAGATGATCTCCCAGTCTTGGTCTGATGGCCGCCGACATGGTGTCTTGTCTCATAGAGCATGGTGATC
 53 I Y A E I H V A S I I E G T K T R I A A G V H H K N D E Y L M T I
 StuI (3105)

3201 TTCTCAGTGGCGACTCCACCAGCTCCAGATCCTGCTGAGAGATGTTGAAGTCTTCATGATGGCCCTCTATAGTGAGTCGTATTATACTATGCCGATA
 19 K E T A V E V L E L D Q S I N F T K M
 BbsI (3251)

3301 TACTATGCCGATGATTAATTGTCAAACACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTTCACATAACGAGCTCTGCTTATATAGACCTCCACCGT
 AseI (3313)

3400 ACACGCCTACCGCCATTGCGTCAATGGGGCGGAGTTGTACGACATTTTGAAAGTCCCGTTGATTACTAGTCAAAAACAACTCCATTGACGTCA
 SpeI (3468)

3499 ATGGGGTGGAGACTTGAAATCCCCGTGAGTCAAACCGCTATCCACGCCATTGATGTACTGCCAAAACCGCATCATCATGGTAATAGCGATGACTAATA
 3599 CGTAGATGTACTGCCAAGTAGGAAAGTCCATAAGGTCATGTACTGGGCATAATGCCAGGCGGGCCATTTACCGTCATTGACGTCAATAGGGGGCGTACT
 SnaBI (3596)

3699 TGGCATATGATACACTTGATGTACTGCCAAGTGGGCGAGTTTACCGTAAATACTCCACCCATTGACGTCAATGGAAAGTCCCTATTGGCGTTACTATGGGA
 NdeI (3701)

3799 ACATACGTCATTATTGACGTCAATGGGCGGGGTCGTTGGGCGGTGAGCCAGGCGGGCCATTTACCGTAAAGTTATGTAACGCCTGCAGGTTAA TTAAG
 SdaI (3879) PacI (3887)

3897 AACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCTTGTGGCTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAA
 BspLU11I (3897)

3997 ATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGGCTCTCCTGTTCCGACCT
 4097 GCCGCTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCATAGCTACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCCG

4197 TCCAAGCTGGGCTGTGTGCACGAACCCCCGTTAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAAGTCCAAACCCGTAAGACACGACTTAT
 4297 CGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGTACAGAGTCTTGAAGTGGTGGCCTAACTACGGCTACACTAG
 4397 AGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGT
 4497 GGTTTTTTGTTTGAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGCTGACGCTCAGTGGAACGAAA

4597 ACTCACGTTAAGGGATTTGGTCATGGCTAGTTAATTAACATTTAAATCAGCGGCCCAATAAAAATATCTTTATTTTATTACATCTGTGTGTTGGTTTT
 Pacl (4627) SwaI (4636) EagI (4647) NotI (4646)

4697 TTGTGTAATCGTAACTAACATACGCTCTCCATCAAAACAAAACGAAACAAAACAACTAGCAAATAGGCTGTCCCGAGTCAAGTGCAGGTGCCAGAA
 4797 CATTCTCTATCGAA