



**PvuI (7)**  
**SgfI (6)** 1 GGATCTGCGATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATCGCCACAGTCCCGGAGAAGTTGGGGGAGGGGTCGGCAATTGAACGGGTGCCTA **EcoNI (96)**

101 GAGAAGGTGGCGCGGGTAAACTGGAAAGTGATGTCGTGTAAGTGGTCCGCCTTTTTCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC

**Psp1406I (203)** 201 **HindIII (245)** **Bsu36I (291)**  
**PvuII (239)** **EcoNI (287)**  
GTGAACGTTCTTTTTTCGCAACGGGTTTCCGCCAGAACACAGCTGAAGCTTCAGAGGGCTCGCATCTCTCTTACCGCCGCCGCCCTACCTGAGGCC

301 GCCATCCACGCGGTTGAGTGCAGTCTGCCGCTCCCGCCTGTGGTGCCTCTGAAGTGCCTCCGCGTCTAGGTAAGTTAAAGCTCAGGTCGAGACC

401 **NgoMIV (441)**  
GGGCTTTGTCCGGCGCTCCCTTGGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCCTGCTTGTCTAACTCTACGCTTTGTTTCGTTT

501 **KasI (535)** **AgeI (552)** **NcoI (560)** **BstEII (555)**  
TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGGGCGCTACCTGAGATCACCGGTCCACATGGCAGTGACATCCCACCACATGGTCCCTGTGTTTGT  
1 M A V T S H H M V P V F V

601 **DraIII (642)**  
CCTGATGAGCGCTGCCTGGCCACCGCAGGTCCAGAGCCAGCACCCGGTGTGAAGTGTACCCGATCAGTGCCTCTCATCCAGTCCAGGCCCTGATGGAG  
13 L M S A C L A T A G P E P S T R C E L S P I S A S H P V Q A L M E

701 AGTTCACCGTTCTGTCTGGCTGTGCCAGCAGAGGCCACTGGGCTGCCAAGGGAGGTTACATCCTAAACCTCCGACAGTACAGACCAAGGACTAGGCC  
47 S F T V L S G C A S R G T T G L P R E V H I L N L R S T D Q G L G

801 **BstEII (813)**  
AGCCGAGAGAGAGGTTACCTGCATCTGAACCCATTGCTCCGTGCACACTCACCACAAGCCTGTTGTCTGCTCAACTCCCCACAGCCCTGGT  
80 Q P Q R E V T L H L N P I A S V H T H H K P V V F L L N S P Q P L V

901 GTGGCATGTGAAGACAGAGAGACTGGCTGCTGGTGTCCCAGACTCTTCTGGTTTCGGAGGGTCTGTGGTCCAGTTTTATCAGGAAACTTCTCCTTG  
113 W H V K T E R L A A G V P R L F L V S E G S V V Q F S S G N F S L

1001 **Bsp120I (1055)**  
ACAGCAGAAACAGAGAAAGGAGTTCCCTCAAGAAAATGAGCATCTGCTACACTGGGCCAAAAGGAATATGGAGCAGTGACTTCATTACCCGAACTTA  
147 T A E T E E R S F P Q E N E H L L H W A Q K E Y G A V T S F T E L

1101 **BbrPI (1152)**  
AGATAGCAAGAAATCTATATTAAGTGGGAGAAGATCAAGTGTCCCTCCACGTAAACATAGGGAAAAATTTCTCTCGCTCAATTACCTTGCGGA  
180 K I A R N I Y I K V G E D Q V F P P T C N I G K N F L S L N Y L A E

1201 GTACCTTCAACCCAAAGCCGCGAAGGTTGTGCTGCGCAGTCAGCCACGAGAAGGAAGTGCATATCATTGAGTTAATCTCCCCAACTCCAATCCT  
213 Y L Q P K A A E G C V L A S Q P H E K E V H I I E L I S P N S N P

1301 **XhoI (1343)** **Bsu36I (1354)**  
TACAGCACCTCCAGGTGGATATAATAATTGACATACGACCTGCTCGAGAGGATCTGAGGTGGTCAAAAACTCGTCTGATCTTGAAGTGCAAAAAAT  
247 Y S T F Q V D I I I D I R P A R E D P E V V K N L V L I L K C K K

1401 CTGCAACTGGGTGATCAAGTCTTTGACGTCAAGGAAAATTGAAAGTTATTGCTCCTGACAGTATTGGCTTTGGAAAAGAGAGTGAACGATCCATGAC  
280 S V N W V I K S F D V K G N L K V I A P D S I G F G K E S E R S M T

1501 AGTGACCAATTTGGTAAGAAATGACATCCCTCCACCAAGAGAATCTGATGAAGTGGCATTGGACAATGGCTACAGCCAGTGACGTCATACACCATA  
313 V T K L V R N D I P S T Q E N L M K W A L D N G Y S P V T S Y T I

1601 **Bsp120I (1698)**  
GCTCCTGTGGCAATAGATTTTATCTTCCGCTA GAGAACAACGAGGAGATGAGAGATGAGGAAGTCCACACCATTCTCCTGAGCTTCGGATCCTGCTGG  
347 A P V A N R F H L R L E N N E E M R D E E V H T I P P E L R I L L

1701 **EcoRV (1780)**  
GCCCTGACCACCTGCCTGCCCTGGACAGCCACCTTCCAAGGGGAAATCCCAATGGAGGTTTCCCTTTCCATTCCCGGATATCCCAGGAGAGGGCTG  
380 G P D H L P A L D S P P F Q G E I P N G G F P F P F P D I P R R G W

1801 GAAGGAGGGAGAAGATAGGATCCCCCGCAAGGAACCCATCATTCCAGAGTTCAATTGCTTCCAGACCACAGAGGCCAGAAAGTGAAGGGGGC  
413 K E G E D R I P R P K E P I I P R V Q L L P D H R E P E E V Q G G

1901 GTGAATATCGCCCTGTCAAGTCAAAATGTGACAAATGAAAGATGGTGTGAGTGTAGACAAAGATTCTTCCAGACCAATGGCTACTCGGGAGTGGAGTCA  
447 V N I A L S V K C D N E K M V V A V D K D S F Q T N G Y S G M E L

2001 **Acc65I (2035)**  
CCCTGTTGGATCCTTCTGCAAAGCCAAGATGAATGGTACCCACTTTGTTCTGGAGTCTCCGCTGAATGGCTGTGGTACTAGACATCGGAGGTCAGCCCC  
480 T L L D P S C K A K M N G T H F V L E S P L N G C G T R H R R S A P

2101 AGATGGTGTGGTTTACTATAACTCTATTGTGGTGCAGGCTCCATCCCCGTTGGGATAGCAGTGGCTGGCCAGACGGCTACGAAGATTTGGAGTCCGGTGT  
513 D G V V Y Y N S I V V Q A P S P G D S S G W P D G Y E D L E S G D

2201 **PstI (2271)** **PvuII (2284)**  
AATGGATTTCTGGAGACACAGATGAAGGAGAACTGCCCCCTGAGCCGTGCTGGAGTGGTAGTGTAACTGCAGCTTGGCGCAGCTGAGGAGTCCCA  
547 N G F P G D T D E G E T A P L S R A G V V V F N C S L R Q L R S P

2301 **XmaI (2379)**  
GTGGCTTCCAGGACCAGCTCGATGGAATGCTACCTTCAATATGGAGCTGTATAACACAGACCTTTTCTGGTGCCTCCCCGGGGTCTTCTCTGTGGC  
580 S G F Q D Q L D G N A T F N M E L Y N T D L F L V P S P G V F S V A

2401 **BglII (2448)**  
AGAGAATGAGCATGTATATGTTGAGGTGTCTGTAAGGCTGACCAAGATCTGGGATTTGCCATCCAACTGCTTTATCTCTCCATACTCAAACCCA  
613 E N E H V Y V E V S V T K A D Q D L G F A I Q T C F I S P Y S N P

2501 GACAGAATGTCTGATTACCCATCATCGAGAACATCTGTCCGAAGGATGACTCTGTGAAGTCTACAGCTCCAAGAGAGTGCACCTCCCCATCCACATG  
647 D R M S D Y T I I E N I C P K D D S V K F Y S S K R V H F P I P H

2601 CTGAAGTGGACAAGAAGCGGTTCCAGCTTTGTGTTCAAGTCCGTGTTCAACACCTCCCTGCTCTTCTGCACTGCGAGCTGACGCTGTGCTAGGAAGAA  
680 A E V D K K R F S F V F K S V F N T S L L F L H C E L T L C S R K K

2701 GGGCTCCCAGAAGCTGCCAAAGTGTGTGACTCCTGATGACGCCTGCACCTCTCTAGATGCCACCATGATCTGGACCATGATGCAGAATAAGAAGACATTC  
713▶ G S Q K L P K C V T P D D A C T S L D A T M I W T M M Q N K K T F  
2801 ACCAAGCCCTGGCCGTGGTCTCCAAGTAGATTATAAAGAAAATGTTCCAAACATGAAGGAGTCCAGTCCGGTTCCTCCTCCACAGATTTTCCACG  
747▶ T K P L A V V L Q V D Y K E N V P N M K E S S P V P P P P Q I F H  
2901 GCCTGGACACGCTCACCGTATGGGCATTGCGTTTGCAGCATTGTGATCGGAGCACTCTGACGGGGCCTTGTGGTATATCTACTCCACACAGGGGA  
780▶ G L D T L T V M G I A F A A F V I G A L L T G A L W Y I Y S H T G E  
3001 GACGGCACGAAGGAGCAAGTCCCTACCTCGCCACCAGCTCGGAGAACAGCAGCGCAGCCACAGCATAGGGAGCACTCAGAGCACTCCCTGCTCTAGC  
813▶ T A R R Q Q V P T S P P A S E N S S A A H S I G S T Q S T P C S S

**XbaI (2751)**

**NheI (3117)**

**AvrII (3110)**

3101 AGCAGCACGGCCTAGGTGCTAGCTGGCCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACCTAGAATGCAGTGAAAAAATGCTTTATTTG  
847▶ S S T

**HpaI (3255)**

3201 TGAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAAACAACAACAAATTGCATTCATTTTATGTTTCAGGTTACAGGG

**EcoRI (3351)**

3301 GAGGTGTGGGAGGTTTTTTAAAGCAAGTAAACCTCTACAATGTGGTATGGAATCTAAAATACAGCATAGCAAACTTTAACCTCCAAATCAAGCCTC  
3401 TACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTGCCAATGTGCATTAGCTGTTGCAGCCTCACCTCTTTTCATGGAGTTT

**SspI (3590)**

3501 AAGATATAGTGATTTTTCCCAAGGTTTGAAGTACTCTTCATTCTTTATGTTTTAAATGCAGTACCTCCACATTCCCTTTTTAGTAAAAATTCAGA

**SwaI (3604)**

3601 AATAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCATAATATCCCCAGTTTAGTAGTTGG

3701 ACTTAGGGAACAAAGAACCTTTAATAGAAATTGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCTGGTGTACTTGAGGGGGATGAGTTCCTCAATG  
141◀ • N R T Y K L P I L E E I

3801 GTGGTTTTGACCAGCTTGCCATTCAATGAGCACAAAGCAGTCCAGGAGCATAGTCCAGAGATGAGCTCTGCACATGCCACAGGGGCTGACCACCC  
127▶ T T K V L K G N 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  
3901 TGATGGATCTGTCCACCTCATCAGAGTAGGGTGCCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCCGTTGCTCACAGCAGACCCAAATGGCAATGGC  
94▶ I S R D V E D S 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

**StuI (4029)**

4001 TTCAGCACAGACAGTGCACCTGCCAATGTAGGCCTCAATGTGGACAGCAGAGATGATCTCCCAAGTCTTGGTCTGATGGCCGCCCGACATGGTGCTTG  
61▶ E A C V T V R G 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

**BspHI (4179)**

4101 TTGCTCATAGAGCATGGTGTCTTCTCAGTGGCAGCTCCACCAGTCCAGATCCTGCTGAGAGATGTTGAAGGTCTTCATGATGGCCCTCTATAGT  
27▶ N D E Y L M T I K E T A V E V L E L D Q Q S I N F T K M

**AseI (4237)**

4201 GAGTCGTATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTTCATAAACGAGCTC

**SpeI (4392)**

4301 TGCTTATATAGACCTCCACCGTACACGCTACCGCCATTTGCGTCAATGGGGCGGAGTTGTTACGACATTTTGAAAGTCCCGTTGATTTACTAGTCA  
4401 AAACAACTCCCATTGACGTCAATGGGGTGGAGACTTGAAATCCCCGTGAGTCAAACCGCTATCCACGCCATTGATGTACTGCCAAAACCGCATCATC

**SnaBI (4520)**

4501 ATGGTAATAGCGATGACTAATACGTAGATGACTGCCAAGTAGGAAAGTCCATAAGGTCACTGACTGGGCATAATGCCAGGCGGGCCATTTACCGTCAT

**NdeI (4625)**

4601 TGACGTCAATAGGGGGCGTACTTGGCATATGATACACTTGATGTACTGCCAAGTGGGCAGTTTACCGTAAATACTCCACCATTGACGTCAATGGAAAGT  
4701 CCCTATTGGCGTTACTATGGGAACATACGTCATTATTGACGTCAATGGGCGGGGCTGTTGGGCGGTGAGCCAGGCGGGCCATTTACCGTAAGTTATGTA

**PacI (4811)**

**PstI (4804)**

**SdaI (4803)**

**BspLU11I (4821)**

4801 ACCGCTGCAGGTTAATTAAGAACATGTGAGCAAAGGCCAGCAAAGGCCAGGAACCGTAAAAAGGCCGCTTGTGGCCTTTTCCATAGGCTCCGCC  
4901 CCCTGACGAGCATCAAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCTGGAAGCTCCCTCGTG  
5001 CGCTCTCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCTTTCTCCCTTGGGAAGCGTGGCGCTTTCTCATAGCTCAGCTGTAGGTATCTCA  
5101 GTTCGGTGTAGGTCGTTGCTCCAAGCTGGGCTGTGTGCACGAACCCCGTTCAGCCCGACCGCTGCGCTTATCCGGTAACTATCGTCTTGTAGTCCAA  
5201 CCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTCTTGAAGTGGTGG  
5301 CCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAAC  
5401 AAACCACCGCTGGTAGCGGTGGTTTTTTGTTTGAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTC

**EagI (5571)**

**PacI (5551)** **SwaI (5560)** **NotI (5570)**

5501 TGACGCTCAGTGGAAACGAAAACCTACGTTAAGGGATTTGGTCATGGCTAGTTAATTAACATTTAAATCAGCGGCCGAATAAAAATATCTTTATTTTCAT  
5601 TACATCTGTGTGTTGTTTTTGTGTGAATCGTAACATAACGCTCTCCATCAAACAAAACGAAACAAAACAACTAGCAAAATAGGCTGTCCCCAG  
5701 TGCAAGTGCAGGTGCCAGAACATTTCTCTATCGAA